

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
15 juillet 2004 (15.07.2004)

PCT

(10) Numéro de publication internationale
WO 2004/058815 A2

(51) Classification internationale des brevets⁷ :
C07K 14/435, C12Q 1/68,
C12N 15/63, C07K 16/18, G01N 33/50, A01K 67/027,
A61K 48/00, 38/00, 39/00, C12N 15/12

Résidence Parc d'Alco, Appt 125, F-34080 Montpellier (FR).

(21) Numéro de la demande internationale :

PCT/FR2003/003895

(74) Mandataire : CABINET ORES; 36 rue de St Pétersbourg, F-75008 Paris (FR).

(22) Date de dépôt international :

24 décembre 2003 (24.12.2003)

(81) États désignés (*national*) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Langue de dépôt :

français

(84) États désignés (*regional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(26) Langue de publication :

français

Publiée :

- sans rapport de recherche internationale, sera republiée dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(30) Données relatives à la priorité :

02/16648 24 décembre 2002 (24.12.2002) FR

(71) Déposant (*pour tous les États désignés sauf US*) : CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE [FR/FR]; 3 rue Michel-Ange, F-75794 Cedex 16 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : GIORGI, Dominique [FR/FR]; 391 rue du Mas du Juge, F-34980 Saint Gely du Fesc (FR). ROUQUIER, Sylvie [FR/FR]; 391 rue du Mas du Juge, F-34980 Saint Gely du Fesc (FR). SAFFIN, Jean-Michel [FR/FR]; 59 rue Michel Teule,

WO 2004/058815 A2

(54) Title: NOVEL CENTROSOME-ASSOCIATED PROTEIN AND APPLICATIONS THEREOF

(54) Titre : NOUVELLE PROTEINE ASSOCIEE AUX CENTROSOMES ET SES APPLICATIONS

(57) Abstract: The invention relates to a novel centrosome-associated protein, to the polynucleotide coding for the aforementioned protein and to the applications of said protein and polynucleotide. The overexpression of the inventive protein disrupts the mitotic spindle assembly and leads to aberrant and abortive mitoses.

(57) Abrégé : L'invention est relative une nouvelle protéine associée aux centrosomes, au polynucléotide codant pour ladite protéine ainsi qu'aux applications de ladite protéine et dudit polynucléotide. La surexpression de la protéine selon l'invention entraîne des perturbations dans l'organisation du fuseau mitotique et induit des mitoses aberrantes et abortives.

THIS PAGE BLANK (USPTO)

NOUVELLE PROTEINE ASSOCIEE AUX CENTROSOMES ET SES APPLICATIONS.

La présente Invention est relative à une nouvelle protéine associée aux centrosomes, au polynucléotide codant pour ladite protéine ainsi 5 qu'aux applications de ladite protéine et dudit polynucléotide.

Le processus de division cellulaire consiste en une division nucléaire (mitose) suivie d'une division cytoplasmique (cytokinèse). La mitose est dominée par la formation d'un fuseau polaire très organisé (le fuseau mitotique) constitué de deux familles de microtubules : les microtubules 10 polaires et les microtubules kinétochoriens. Les microtubules sont des polymères composés de sous-unités d' α - et β -tubuline. Leur croissance est initiée dans la région périphérique du centrosome par un complexe contenant majoritairement une protéine apparentée, la γ -tubuline. Les microtubules polaires sont composés de rangées de microtubules et de protéines associées qui 15 sont mises en place par les deux centres mitotiques, associés à des centrioles, situés aux pôles opposés du fuseau (asters). Chaque chromosome répliqué est constitué de deux chromatides sœurs reliées entre elles par le centromère. Les microtubules kinétochoriens sont liés aux chromosomes répliqués par des structures spécialisées appelées kinétochores qui se 20 forment au cours de la prophase sur chacune des deux faces du centromère. Les chromosomes se condensent pendant la prophase et forment les microtubules kinétochoriens qui commencent à interagir avec les microtubules polaires du fuseau après rupture de l'enveloppe nucléaire au cours de la prométaphase. Sous l'effet de la tension due aux forces opposées, dirigées 25 vers les pôles qui tirent les microtubules kinétochoriens, les chromosomes s'alignent dans la zone équatoriale du fuseau pendant la métaphase. A l'anaphase, sous l'effet de forces continuellement développées au sein du fuseau mitotique, les chromatides sœurs se détachent et sont attirées vers les pôles opposés. Dans le même temps, les deux pôles cellulaires s'écartent. Au cours 30 de la télophase, l'enveloppe nucléaire se reforme à la surface de chaque groupe de chromosomes.

La division cellulaire s'achève au moment où le contenu cytoplasmique est divisé selon le processus de cytokinèse. Le fuseau mitotique joue un rôle important dans le processus de cytokinèse, en fixant la mise en place de la segmentation cellulaire. Le sillon de division apparaît invariablement dans le plan de la plaque équatoriale, perpendiculairement à l'axe du fuseau mitotique.

Les processus décrits ci-dessus sont finement régulés par un équilibre entre des réactions de phosphorylation et de déphosphorylation. Lorsque la cellule entre en mitose, des changements importants dans la phosphorylation des protéines interviennent. Le centrosome et le fuseau mitotique sont particulièrement enrichis en sites phosphorylés. De nombreuses protéine-kinases, particulièrement des sérine-thréonine-kinases, ont été décrites comme intervenant dans ces processus de phosphorylation (voir à cet égard Giet R. et Prigent C., J. Cell Science, 112, 3591-3601, 1999).
Parmi celles-ci on citera celles localisées au niveau des centrosomes, parmi lesquelles les kinases de type aurora, requises pour la séparation des centrosomes et l'assemblage du fuseau mitotique, les kinases de type polo, impliquées dans la maturation et la formation du fuseau bipolaire et les kinases de type NIMA qui régulent la séparation des centrosomes.

Les mammifères possèdent au moins trois protéine-kinases du type aurora. Chez l'homme, ces trois protéine-kinases sont surexprimées dans des pathologies cancéreuses du fait d'anomalies chromosomiques. Ainsi, ces protéines semblent jouer un rôle important dans le contrôle de la ploïdie. Par exemple, une inactivation ou une surexpression de deux de ces kinases conduit à une polyploïdie. L'inhibition de l'activité de la kinase aurora A conduit à la formation de fuseaux monopolaires. L'inhibition de l'activité de la kinase aurora B conduit à la formation de cellules multinucléées par défaut de cytokinèse. Ces anomalies chromosomiques apparaissent liées à des perturbations dans la formation du fuseau mitotique.

Les partenaires et les substrats de ces protéine-kinases sont encore peu connus. Par exemple, chez le xénope, aurora A interagit avec une kinésine impliquée dans la dynamique des microtubules. Chez l'homme, elle

phosphoryle la protéine HsTACC-3, également surexprimée dans de nombreuses lignées de cellules cancéreuses. Chez la drosophile, aurora A phosphoryle la protéine D-TACC et est nécessaire à sa localisation aux centrosomes afin de réguler les microtubules astraux. D-TACC interagit avec 5 la protéine associée aux microtubules (MAP : Microtubule Associated Protein) Msp, qui fait partie de la famille des protéines XMAO215/ch-TOC/Msp, qui stimulent la croissance des microtubules *in vitro* et sont concentrées au niveau des centrosomes *in vivo*. D-TACC et Msp coopèrent pour stabiliser les centrosomes. Le terme MAP regroupe une collection de protéines variées 10 définies sur la base de leur capacité à interagir avec les microtubules. Les MAP apparaissent comme des partenaires/substrats des kinases du centrosome comme aurora ou polo.

Une division cellulaire correcte nécessite une coordination entre la ségrégation des chromosomes par le fuseau mitotique et le clivage de 15 la cellule par l'appareil de cytokinèse. Les microtubules du fuseau mitotique jouent un rôle essentiel dans les deux processus.

Cependant, malgré l'ensemble des travaux réalisés sur la division cellulaire, les facteurs intervenant dans une mise en place correcte du fuseau mitotique et/ou au contraire perturbant sa mise en place et/ou sa 20 structure, entraînant ainsi les conséquences ci-dessus décrites ne sont toujours pas connus.

Une telle connaissance permettrait d'une part de mieux comprendre les mécanismes de la mitose et d'autre part de pouvoir développer des moyens de lutter contre les anomalies de la division cellulaire et les 25 conséquences qu'elles entraînent.

C'est dans ce domaine que se place la présente Invention.

En effet, de manière surprenante et inattendue, les Inventeurs ont mis en évidence une nouvelle protéine humaine associée aux centrosomes. Par immunofluorescence, elle est détectée en colocalisation avec l' α -tubuline des microtubules du fuseau mitotique, en particulier avec l'aster. 30 Cette protéine a été nommée ASAP pour Aster Associated Protein (Protéine Associée à l'Aster) par les Inventeurs.

La surexpression de la protéine selon l'Invention entraîne des perturbations dans l'organisation du fuseau mitotique et induit des mitoses aberrantes et abortives (cellules plurinucléées, fuseaux mono- ou multipolaires). Sa surexpression bloque la division cellulaire et par conséquent la 5 prolifération cellulaire.

Ainsi, l'Invention a pour objet une protéine isolée, dénommée ASAP, caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par :

- a) une protéine répondant à la séquence représentée dans la 10 liste de séquences en annexe sous le numéro SEQ ID NO: 1 ;
- b) une protéine présentant, sur la totalité de sa séquence, au moins 80% d'identité ou au moins 90% de similarité, de préférence au moins 90 % d'identité ou au moins 95% de similarité, avec la protéine de SEQ ID NO: 1.

15 Une protéine conforme à l'Invention se caractérise par les propriétés suivantes :

- elle présente un poids moléculaire compris entre 60 et 100 kDa, de préférence entre 65 et 80 kDa ;
- elle est associée aux centrosomes ;
- elle est colocalisée par immunofluorescence avec l'α-tubuline des microtubules du fuseau mitotique ;
- elle présente une faible identité (23%) avec la protéine MAP1A (Microtubule Associated Protein 1A) ;
- elle présente des domaines coiled-coil essentiellement regroupés dans sa partie C-terminale entre les acides aminés 297 et 327 d'une part et 477 et 628 d'autre part, indiquant soit que la protéine s'oligomérisé, soit qu'elle interagit avec d'autres protéines ;
- elle présente une faible identité (20%), entre les acides aminés 300 et 600 avec un domaine de type caldesmon (Gusev, N.B., Biochemistry, 10 : 1112-1121, 2000), référencé pfam00769 (NCBI, domains, 20 <http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrvcgi?uid=pfam00769>), et, entre les acides aminés 480 et 630, avec un domaine de type ERM .

(Ezrin/radixin/moesin ; Louvet-Vallet, S., Biol. Cell, 274 : 305-316, 2000), référencé pfam02029 (NCBI, domains, <http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrvcgi?uid=pfam02029>). Les protéines caldesmon et ERM sont également considérées comme des MAP ;

- 5 - elle présente également, entre les positions 65 et 303, un domaine BRCT, (Breast cancer carboxy-terminal domain ; Bork, P., et coll., FASEB J., 11, 68-76 (1997)), indiquant que la protéine est impliquée dans le contrôle du cycle cellulaire ;
- 10 - elle présente une grande richesse en hélices α dans sa partie C-terminale, en particulier dans la région comprise entre les acides aminés 420-620, presque exclusivement formée d'hélices α .

Ces éléments permettent de penser que la protéine ASAP est une nouvelle MAP.

15 Les protéines selon l'Invention incluent toute protéine (naturelle, synthétique, semi-synthétique ou recombinante) de n'importe quel organisme procaryote ou eucaryote, notamment d'un mammifère, comprenant ou consistant en une protéine ASAP. Préférentiellement, ladite protéine est une protéine ASAP fonctionnelle.

20 On entend par "fonctionnelle", une protéine possédant une activité biologique normale, c'est à dire capable d'intervenir dans l'organisation du fuseau mitotique et dans la division cellulaire. Cette protéine peut comprendre des mutations silencieuses n'induisant aucun changement substantiel dans son activité et ne produisant aucune modification phénotypique.

25 Des protéines conformes à l'invention sont notamment représentées par les protéines ASAP humaine (SEQ ID NO : 1) et murine (SEQ ID NO : 46).

30 Sont incluses dans les protéines selon l'Invention définies en b), les protéines variantes des séquences SEQ ID NO: 1 et 46, en particulier les protéines dont la séquence en acides aminés présente au moins une mutation correspondant notamment à une troncation, une délétion, une substitution et/ou une addition d'au moins un résidu d'acide aminé par rapport

aux séquences SEQ ID NO: 1 et 46.

De manière préférée, les protéines variantes présentent une mutation entraînant un dysfonctionnement (activation ou inhibition) de la protéine, d'autres gènes ou protéines ou encore de la cellule en général.

5 Selon un autre mode de réalisation avantageux de l'Invention, ladite protéine est une protéine de mammifère, préférentiellement une protéine d'origine humaine.

Au sens de la présente Invention les définitions suivantes s'appliquent.

10 L'identité d'une séquence par rapport à la séquence de SEQ ID NO :1 comme séquence de référence s'apprécie en fonction du pourcentage de résidus d'acides aminés qui sont identiques, lorsque les deux séquences sont alignées, de manière à obtenir le maximum de correspondance entre elles.

15 Le pourcentage d'identité peut être calculé par l'Homme du métier en utilisant un programme informatique de comparaison de séquences tel que, par exemple celui de la suite BLAST (Altschul et al., NAR, 1997, 25, 3389-3402).

20 Les programmes BLAST sont mis en œuvre sur la fenêtre de comparaison constituée par la totalité de la SEQ ID NO :1, indiquée comme séquence de référence.

25 Une protéine ayant une séquence en acides aminés ayant au moins X % d'identité avec une séquence de référence est définie, dans la présente Invention comme une protéine dont la séquence peut inclure jusqu'à 100-X altérations pour 100 acides aminés de la séquence de référence, tout en conservant les propriétés fonctionnelles de ladite protéine de référence. Au sens de la présente Invention, le terme altération inclut les délétions, les substitutions ou les insertions consécutives ou dispersées d'acides aminés dans la séquence de référence.

30 La similarité d'une séquence par rapport à une séquence de référence s'apprécie en fonction du pourcentage de résidus d'acides aminés qui sont identiques ou qui diffèrent par des substitutions conservatives,

lorsque les deux séquences sont alignées de manière à obtenir le maximum de correspondance entre elles. Au sens de la présente Invention, on entend par substitution conservative, la substitution d'un acide aminé par un autre qui présente des propriétés chimiques ou physiques similaires (taille, charge ou 5 polarité), qui généralement ne modifie pas les propriétés fonctionnelles de la protéine.

Une protéine ayant une séquence en acides aminés ayant au moins X % de similarité avec une séquence de référence est définie, dans la présente Invention comme une protéine dont la séquence peut inclure jusqu'à 10 100-X altérations non-conservatives pour 100 acides aminés de la séquence de référence. Au sens de la présente Invention, le terme altérations non-conservatives inclut les délétions, les substitutions non-conservatives ou les insertions consécutives ou dispersées d'acides aminés dans la séquence de référence.

15 Par "techniques ou méthodes bien connues de l'homme du métier" on entend ici se référer aux techniques ou méthodes classiquement utilisées par l'homme du métier et exposées dans de nombreux ouvrages, comme en particulier celui intitulé Molecular Cloning. A Laboratory Manual (Sambrook J, Russell DW. (2000) Cold Spring Harbor Laboratory Press).

20 La protéine selon l'Invention est obtenue soit à partir d'une cellule, soit par synthèse chimique, soit par recombinaison génétique.

Par synthèse chimique, la protéine peut être obtenue en utilisant l'une des nombreuses voies de synthèses peptidiques connues, par exemple les techniques mettant en œuvre des phases solides ou des 25 techniques utilisant des phases solides partielles, par condensation de fragments ou par une synthèse en solution classique. Dans ce cas, la séquence de la protéine peut être modifiée afin d'améliorer sa solubilité, en particulier dans les solvants aqueux. De telles modifications sont connues de l'homme du métier comme par exemple la délétion de domaines hydrophobes ou la 30 substitution d'acides aminés hydrophobes par des acides aminés hydrophiles.

La protéine selon l'Invention est constituée de l'enchaînement de 13 peptides correspondants aux produits de traduction de 13 des 14 exons

que comporte le gène correspondant, le premier exon n'étant pas traduit (voir ci-après).

De manière plus précise, lesdits peptides répondent aux séquences suivantes (positions données par rapport à la numérotation de la 5 séquence SEQ ID NO: 1) :

- Peptide 1 : il comprend 25 acides aminés correspondants aux positions 1 à 25 (SEQ ID NO: 2) ;

- Peptide 2 : il comprend 28 acides aminés correspondants aux positions 26 à 53 (SEQ ID NO: 3) ;

10 - Peptide 3 : il comprend 107 acides aminés correspondants aux positions 54 à 160 (SEQ ID NO: 4) ;

- Peptide 4 : il comprend 76 acides aminés correspondants aux positions 161 à 236 (SEQ ID NO: 5) ;

15 - Peptide 5 : il comprend 31 acides aminés correspondants aux positions 237 à 267 (SEQ ID NO: 6) ;

- Peptide 6 : il comprend 83 acides aminés correspondants aux positions 268 à 350 (SEQ ID NO: 7) ;

- Peptide 7 : il comprend 24 acides aminés correspondants aux positions 351 à 374 (SEQ ID NO: 8) ;

20 - Peptide 8 : il comprend 54 acides aminés correspondants aux positions 375 à 428 (SEQ ID NO: 9) ;

- Peptide 9 : il comprend 32 acides aminés correspondants aux positions 429 à 460 (SEQ ID NO: 10) ;

25 - Peptide 10 : il comprend 54 acides aminés correspondants aux positions 461 à 514 (SEQ ID NO: 11) ;

- Peptide 11 : il comprend 49 acides aminés correspondants aux positions 515 à 563 (SEQ ID NO: 12) ;

- Peptide 12 : il comprend 43 acides aminés correspondants aux positions 564 à 606 (SEQ ID NO: 13) ;

30 - Peptide 13 : il comprend 41 acides aminés correspondants aux positions 607 à 647 (SEQ ID NO: 14).

La présente Invention a aussi pour objet un peptide constitué d'un fragment d'au moins 10 acides aminés consécutifs d'une protéine définie ci-dessus en a) ou b), particulièrement un peptide sélectionné parmi :

5 - les séquences correspondant aux peptides 1 à 13 décrits ci-dessus, c'est-à-dire sélectionné parmi les séquences SEQ ID NO: 2 à SEQ ID NO: 14, et

10 - les séquences SEQ ID NO : 47 à 53 correspondant à des mutants de la protéine hASAP délétés de la partie N-terminale contenant le domaine BRCT (Ndel1 : résidus 304 -647 (SEQ ID NO: 48) ; Ndel2 : résidus 411-647 (SEQ ID NO : 49) ; Ndel3 : résidus 478-647 (SEQ ID NO : 50)) ou de la partie C-terminale contenant le domaine MAP (Cdel1 : résidus 1 à 477 (SEQ ID NO: 51) ; Cdel2 : résidus 1 à 418 (SEQ ID NO: 52) ; Cdel3 : résidus 1 à 303 (SEQ ID NO: 53) ; résidus 1 à 421 (SEQ ID NO: 47)).

15 Selon un mode de réalisation avantageux de l'Invention, ledit peptide est utile pour la production d'anticorps reconnaissant spécifiquement une protéine telle que définie ci-dessus, préférentiellement reconnaissant la protéine ASAP de séquence SEQ ID NO: 1 ou SEQ ID NO: 46.

20 L'Invention a ainsi également pour objet des anticorps monoclonaux ou polyclonaux, caractérisés en ce qu'ils sont capables de reconnaître spécifiquement une protéine selon l'Invention.

De manière préférentielle selon l'Invention, les anticorps reconnaissent, parmi les MAPs, uniquement et spécifiquement la protéine ASAP de séquence SEQ ID NO: 1 ou SEQ ID NO: 46.

25 Les anticorps selon l'Invention sont, par exemple, des anticorps chimériques, des anticorps humanisés, des fragments Fab ou F(ab')2. Ils peuvent également se présenter sous forme d'immunoconjugués ou d'anticorps marqués afin d'obtenir un signal délectable et/ou quantifiable.

30 Lesdits anticorps peuvent être obtenus directement à partir de sérum humain ou à partir de sérum d'animaux immunisés avec les protéines ou les peptides selon l'Invention. Les anticorps polyclonaux ou monoclonaux spécifiques peuvent être obtenus selon les techniques bien connues de l'Homme du Métier.

L'Invention a également pour objet l'utilisation des anticorps selon l'Invention pour la détection et/ou la purification d'une protéine selon l'Invention.

De manière générale, les anticorps selon l'Invention peuvent être avantageusement utilisés pour détecter la présence d'une protéine selon l'Invention, normale ou mutée.

Particulièrement, les anticorps monoclonaux, peuvent être utilisés pour la détection de ces protéines dans un échantillon biologique. Ils constituent ainsi un moyen d'analyse immunocytochimique ou immuno-histochimique de l'expression des protéines selon l'Invention, notamment la protéine de séquence SEQ ID NO: 1, sur des coupes de tissus. Généralement pour de telles analyses, les anticorps utilisés sont marqués afin d'être détectables par exemple par des composés immunofluorescents, par marquage à l'or ou sous forme d'immunoconjugués enzymatiques.

Ils peuvent permettre notamment de mettre en évidence une expression anormale de ces protéines dans les tissus ou prélèvements biologiques et ainsi permettre la détection de cellules présentant des perturbations dans l'organisation du fuseau mitotique et/ou une induction des mitoses aberrantes et abortives (cellules plurinucléées, fuseaux mono- ou multipolaires) liées à la surexpression de la protéine selon l'Invention.

L'Invention a également pour objet une méthode de détection dans un échantillon biologique de la protéine selon l'Invention, particulièrement de la protéine ASAP, comprenant une première étape de traitement convenable des cellules par tout moyen approprié permettant de rendre accessible le milieu intracellulaire, une seconde étape de mise en contact dudit milieu intracellulaire ainsi obtenu avec un anticorps selon l'Invention et une troisième étape de mise en évidence par tout moyen approprié du complexe protéine ASAP-anticorps formé.

Cette méthode peut en outre permettre de mesurer le taux d'expression de la protéine selon l'Invention dans des cellules, particulièrement dans des cellules cancéreuses. L'étude de l'expression de la protéine ASAP (sur- ou sous-expression) est un élément d'évaluation de la capacité de

prolifération ou d'agressivité (capacité à évoluer vers des cancers de mauvais pronostic) de cellules cancéreuses.

L'Invention a donc également pour objet une méthode d'évaluation *in vitro* de la capacité de prolifération ou d'agressivité des cellules cancéreuses contenues dans un échantillon biologique, caractérisée en ce qu'elle comprend une première étape de traitement convenable des cellules par tout moyen approprié permettant de rendre accessible le milieu intracellulaire, une seconde étape de mise en contact dudit milieu intracellulaire ainsi obtenu avec un anticorps selon l'Invention, une troisième étape de mise en évidence et/ou de mesure par tout moyen approprié du complexe protéine ASAP-anticorps formé et une quatrième étape d'évaluation du taux de transcription du gène par comparaison du taux de complexes protéine ASAP-anticorps formés à celui d'un échantillon biologique témoin préalablement choisi. Ledit témoin peut être constitué par exemple par un échantillon biologique contenant des cellules présentant un taux de protéines normal ou altéré, auquel ladite méthode est appliquée dans les mêmes conditions.

L'Invention a également pour objet une trousse permettant de mettre en œuvre l'une quelconque des méthodes ci-dessus décrites comprenant :

- 20 a) au moins un anticorps monoclonal ou polyclonal selon l'Invention ;
- b) les réactifs permettant la détection du complexe protéine ASAP-anticorps produit lors de la réaction immunologique.

Selon un mode de réalisation particulier de l'Invention, la trousse peut éventuellement comprendre des réactifs nécessaires pour rendre accessible le milieu intracellulaire.

Par moyen pour rendre accessible le milieu intracellulaire, on entend tout moyen connu de l'Homme du Métier comme par exemple la lyse cellulaire par voie enzymatique, chimique ou encore la sonication, la perméation membranaire, les chocs thermiques.

La présente Invention a également pour objet un polynucléotide isolé (ADNc ou fragment d'ADN génomique), caractérisé en ce que sa séquence est sélectionnée dans le groupe constitué par :

- les séquences codant pour une protéine ou un peptide
- 5 tels que définis ci-dessus, et
 - les séquences complémentaires des précédentes, sens ou anti-sens.

L'Invention englobe, les allèles du gène *asap* issus de n'importe quel mammifère, ainsi que les polynucléotides des mutants naturels 10 ou artificiels du gène *asap* codant pour une protéine ASAP, particulièrement pour une protéine ASAP fonctionnelle telle que définie ci-dessus.

Selon un mode de réalisation avantageux de l'Invention, ledit polynucléotide codant pour une protéine ASAP répond à une séquence sélectionnée dans le groupe constitué par :

- 15 - la séquence SEQ ID NO: 15, correspondant à l'ADN complémentaire de 2575 nucléotides de l'ARNm codant pour la protéine ASAP humaine (hASAP) ;
 - la séquence SEQ ID NO: 45, correspondant à l'ADN complémentaire de 2767 nucléotides de l'ARNm codant pour la protéine 20 ASAP murine (mASAP)

25 - le fragment d'ADN génomique de 29750 nucléotides répondant à la séquence représentée dans la liste des séquences en annexe sous le numéro SEQ ID NO: 16, correspondant au gène *asap* humain comprenant 14 exons dont 13 seulement sont traduits, le premier exon n'étant pas traduit, contenue dans le contig AC097467 (longueur 178204 paires de bases) entre les bases 115117 et 143828 (version v.7.29a3 NCBI/Ensembl du 12 juillet 2002, <http://www.ensembl.org>), par ailleurs localisée sur le chromosome 4q32.1 entre les marqueurs anonymes D4S1053 et D4S571 (région 161,25 Mégabases (Mb) à 161,28 Mb).

30 La séquence SEQ ID NO: 16 est contenue dans le clone BAC RP11-27G13 (Osoegawa, K., et col., (2001) A Bacterial Artificial Chromosome Library for Sequencing the Complete Human Genome, Genome Research,

Vol. 11, n°3, 483-496, mars 2001). Les séquences contenues dans le contig AC097467 et dans le clone BAC RP11-27G13 ont été obtenues dans le cadre du programme de séquençage du génome humain et n'ont jusqu'à présent fait l'objet d'aucune reconnaissance ni caractérisation précises permettant de leur attribuer une quelconque fonction. Deux acides nucléiques correspondants à des fragments du polynucléotide isolé par les Inventeurs sont répertoriés dans la base de données GenBank sous les numéros d'accès AK024730 et AK024812, ainsi que les EST répertoriés sous les numéros d'accès BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952. Ces séquences, obtenues dans le cadre d'un programme de séquençage en masse de banques d'ADN complémentaires humains, sont incomplètes et n'ont jamais été ni reconnues ni caractérisées. En fait, le polynucléotide isolé par les Inventeurs présente de longs enchaînements de désoxyadénosines (poly-dA), ce qui explique les difficultés rencontrées par les Inventeurs pour obtenir l'ADNc complet par utilisation d'amorces oligo-désoxythymidines (oligo-dT) classiques, celles-ci s'hybridant

de manière aléatoire avec les enchaînements poly-dA. C'est par l'utilisation répétée de la technique de l'amplification rapide des extrémités 3' d'ADNc (3' Rapid Amplification cDNA end ou 3'RACE) que les Inventeurs sont parvenus à isoler le polynucléotide correspondant à l'ARNm complet.

5 L'ARNm, correspondant au polynucléotide de séquence SEQ ID NO: 15, est spécifiquement exprimé dans le testicule sous la forme d'un polynucléotide d'une longueur d'environ 2,9 kilobases et dans le cerveau sous la forme d'un polynucléotide d'une longueur d'environ 9 kilobases pouvant correspondre soit à un prémessager soit à une isoforme de haut poids moléculaire.

10 De manière plus précise, lesdits exons sont répartis comme suit sur ladite séquence génomique (par rapport à la numérotation de la séquence SEQ ID NO: 16) :

- exon 1 : il comprend 200 paires de bases correspondant aux positions 101 à 300 (SEQ ID NO: 17) ;
- exon 2 : il comprend 139 paires de bases correspondant aux positions 1157 à 1295 (SEQ ID NO: 18) ;
- exon 3 : il comprend 85 paires de bases correspondant aux positions 2050 à 2134 (SEQ ID NO: 19) ;
- 15 - exon 4 : il comprend 321 paires de bases correspondant aux positions 3615 à 3935 (SEQ ID NO: 20) ;
- exon 5 : il comprend 227 paires de bases correspondant aux positions 8259 à 8485 (SEQ ID NO: 21) ;
- exon 6 : il comprend 94 paires de bases correspondant aux positions 14930 à 15023 (SEQ ID NO: 22) ;
- exon 7 : il comprend 248 paires de bases correspondant aux positions 16715 à 16962 (SEQ ID NO: 23) ;
- exon 8 : il comprend 71 paires de bases correspondant aux positions 19552 à 19622 (SEQ ID NO: 24) ;
- 20 - exon 9 : il comprend 169 paires de bases correspondant aux positions 21187 à 21355 (SEQ ID NO: 25) ;

- exon 10 : il comprend 90 paires de bases correspondant aux positions 21911 à 22000 (SEQ ID NO: 26) ;
 - exon 11 : il comprend 162 paires de bases correspondant aux positions 23731 à 23892 (SEQ ID NO: 27) ;
- 5 - exon 12 : il comprend 146 paires de bases correspondant aux positions 24014 à 24159 (SEQ ID NO: 28) ;
- exon 13 : il comprend 133 paires de bases correspondant aux positions 24343 à 24475 (SEQ ID NO: 29) ;
 - exon 14 : il comprend 485 paires de bases correspondant aux
- 10 positions 29166 à 29650 (SEQ ID NO: 30).

L'Invention a aussi pour objet :

- un fragment de l'un quelconque des polynucléotides selon l'Invention, d'au moins 15 à 1500 nucléotides consécutifs à l'exclusion des séquences répertoriées sous les numéros d'accès AK024730 et AK024812 et 15 des EST répertoriés sous les numéros d'accès BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, 20 AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, 25 BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, 30 BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952 dans

la base de données GenBank, particulièrement un fragment sélectionné parmi les séquences correspondants aux exons c'est-à-dire sélectionné parmi les séquences SEQ ID NO: 16 à SEQ ID NO: 30 ;

- un acide nucléique présentant un pourcentage d'identité 5 d'au moins 80 %, de préférence d'au moins 90 %, avec l'un des poly-nucléotides selon l'Invention.

La définition de l'identité d'une séquence donnée précédemment pour les protéines, s'applique par analogie aux molécules d'acide nucléique.

10 Sont inclus dans un polynucléotide présentant un pourcentage d'identité d'au moins 80 %, de préférence au moins 90 %, selon l'Invention, les polynucléotides variants des séquences SEQ ID NO: 15 et 45, c'est-à-dire l'ensemble des polynucléotides correspondants à des variants alléliques, c'est-à-dire à des variations individuelles des séquences SEQ ID NO: 15 et 15 45. Ces séquences variantes naturelles correspondent à des polymorphismes présents chez les mammifères, en particulier chez l'être humain et, notamment à des polymorphismes pouvant conduire à la survenue d'une pathologie.

On entend également désigner par polynucléotide variant, tout ARN ou ADNc résultant d'une mutation et/ou d'une variation d'un site 20 d'épissage de la séquence génomique dont l'ARNm a comme ADN complémentaire le polynucléotide de séquence SEQ ID NO: 15 ou SEQ ID NO : 45.

De préférence, la présente Invention concerne les polynucléotides ou les fragments variants des séquences SEQ ID NO: 15 et 45, particulièrement ceux dans lesquelles les mutations conduisent à une modification 25 de la séquence en acides aminés des protéines de séquence SEQ ID NO: 1 et SEQ ID NO : 46.

Les polynucléotides selon l'Invention peuvent être isolés à partir de cellules, particulièrement des cellules de testicule ou de cerveau ou à partir de banques d'ADN cellulaire. Ils peuvent également être obtenus par 30 une réaction de polymérisation en chaîne (PCR) effectuée sur l'ADN total des cellules ou encore par RT-PCR effectuée sur les ARN totaux des cellules ou par synthèse chimique.

Les polynucléotides selon l'Invention, particulièrement les fragments de l'un quelconque des polynucléotides selon l'Invention, et les séquences répertoriées sous les numéros d'accès AK024730 et AK024812 et les EST répertoriés sous les numéros d'accès BU198882, BM693711,
5 AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494,
10 BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289,
15 BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559,
20 BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952 dans la base de données GenBank ou leurs fragments, peuvent notamment être utilisés comme sondes ou comme amorces pour détecter/amplifier des polynucléotides (ARN ou ADN génomique) correspondants au polynucléotide
25 selon l'Invention, particulièrement dans d'autres organismes.

Les transcrits du gène *asap* sont par exemple de préférence mis en évidence à l'aide de sondes sélectionnées dans le groupe constitué par les séquences SEQ ID NO: 15, SEQ ID NO : 45, SEQ ID NO: 17 à SEQ ID NO: 44 ou à l'aide d'un EST tel que défini ci-dessus ou amplifiés par RT-PCR à l'aide d'amorces sélectionnées dans le groupe constitué par les séquences SEQ ID NO: 31 à 43.

Le polynucléotide selon l'Invention peut permettre de diagnostiquer un état pathologique ou une maladie génétique impliquant un dysfonctionnement du gène *asap* et de cribler des substances capables de moduler (activer ou inhiber) la transcription dudit gène.

5 L'Invention a aussi pour objet les polynucléotides susceptibles d'être obtenus par amplification à l'aide des amorces selon l'Invention.

Les sondes et amorces selon l'Invention peuvent être marquées directement ou indirectement par un composé radioactif ou non radioactif par des méthodes bien connues de l'homme du métier, afin d'obtenir 10 un signal délectable et/ou quantifiable.

Le marquage des sondes selon l'Invention est réalisé par des éléments radioactifs ou par des molécules non radioactives. Parmi les isotopes radioactifs utilisés, on peut citer le ^{32}P , le ^{33}P , le ^{35}S , le ^3H ou l' ^{125}I . Les entités non radioactives sont sélectionnées parmi les ligands tels que la biotine, l'avidine, la streptavidine, la digoxygénine, les haptènes, les colorants, les agents luminescents tels que les agents radioluminescents, chémo-luminescents, bioluminescents, fluorescents, phosphorescents.

Les polynucléotides selon l'Invention peuvent ainsi être utilisés comme amorce et/ou sonde dans des procédés mettant en œuvre 20 notamment la technique de PCR (amplification en chaîne par polymérase)

(U.S. N° 4,683,202). D'autres techniques d'amplification de l'acide nucléique cible peuvent être avantageusement employées comme alternative à la PCR. Il existe actuellement de très nombreux procédés permettant cette amplification, comme par exemple la technique SDA (Strand Displacement Amplification) ou technique d'amplification à déplacement de brin, la technique TAS

(Transcription-based Amplification System), la technique 3SR (Self-Sustained Sequence Replication), la technique NASBA (Nucleic Acid Sequence Based Amplification), la technique TMA (Transcription Mediated Amplification), la technique LCR (Ligase Chain Reaction), la technique de RCR (Repair Chain

Reaction), la technique CPR (Cycling Probe Reaction), la technique d'amplification à la Q-βeta-réplicase. On peut encore citer la PCR-SSCP qui permet de détecter des mutations ponctuelles.

Ces techniques sont bien entendu parfaitement connues de l'homme du métier.

Comme sondes ou comme amorces, les différents poly-nucléotides selon l'Invention peuvent permettent, soit de déterminer le profil 5 de transcription du gène asap correspondant ou une éventuelle altération de ce profil dans un échantillon biologique, soit de mettre en évidence le gène correspondant dans d'autres espèces, des variants alléliques de ce gène ou une éventuelle altération fonctionnelle de ce gène (changement substantiel dans l'activité de la protéine codée par ledit gène) résultant d'une mutation 10 (insertion, délétion ou substitution) d'un ou plusieurs nucléotides au niveau d'au moins un exon dudit gène. De telles mutations incluent en particulier les délétions, les insertions ou les substitutions non-conservatives au niveau de codons correspondant à des résidus d'acides aminés situés dans un domaine essentiel pour l'activité biologique de la protéine.

15 Ainsi l'Invention a pour objet une méthode de détermination du profil de transcription du gène correspondant au polynucléotide selon l'Invention ou d'une altération dudit profil, dans un échantillon biologique, comprenant une première étape d'obtention par tout moyen approprié des ARN totaux à partir de l'échantillon biologique, une deuxième étape de mise 20 en contact desdits ARN avec une sonde selon l'Invention, préalablement marquée, dans des conditions classiques d'hybridation entre les ARN et la sonde et une troisième étape de révélation par tout moyen approprié des hybrides formés.

Par conditions classiques d'hybridation, on entend celles 25 décrites dans Sambrook J, Russell DW. (2000) Cold Spring Harbor Laboratory Press.

Selon un mode de mise en œuvre de ladite méthode, la 30 deuxième étape peut être une étape de transcription inverse et d'amplification des transcrits, réalisée à l'aide d'une paire d'amorces telles que décrites précédemment et la troisième étape, une étape de révélation par tout moyen approprié des acides nucléiques amplifiés formés.

Ladite méthode de détermination du profil de transcription du gène peut en outre comporter une étape d'évaluation du taux de transcription du gène par comparaison avec un échantillon témoin préalablement choisi. Ledit témoin peut être constitué par exemple par un échantillon biologique 5 présentant une transcription normale ou altérée du gène correspondant au polynucléotide selon l'Invention auquel ladite méthode de détermination du profil de transcription du gène est appliquée dans les mêmes conditions.

L'Invention a aussi pour objet une méthode de mise en évidence dans d'autres espèces du gène correspondant au polynucléotide 10 selon l'Invention ou des variants alléliques dudit gène ou d'une altération fonctionnelle de ce gène, dans un échantillon biologique, comprenant une première étape d'obtention par tout moyen approprié de l'ADN à partir des cellules d'un échantillon biologique, une deuxième étape de mise en contact desdits ADN avec une sonde selon l'Invention, préalablement marquée, dans 15 des conditions classiques d'hybridation entre les ADN et la sonde et une troisième étape de révélation par tout moyen approprié des hybrides formés.

Selon un mode de mise en œuvre de ladite méthode, la deuxième étape peut être une étape d'amplification réalisée à l'aide d'une paire d'amorces telles que décrites précédemment et la troisième étape, une 20 étape de révélation par tout moyen approprié des acides nucléiques amplifiés formés. La méthode peut éventuellement comporter une quatrième étape d'isolement et de séquençage des acides nucléiques mis en évidence.

L'Invention a également pour objet une trousse de réactifs pour la mise en œuvre des méthodes précédemment décrites comprenant :
25 a) au moins une sonde ou une paire d'amorces selon l'Invention ;

b) les réactifs nécessaires à la mise en œuvre d'une réaction classique d'hybridation entre ladite sonde ou lesdites amorces et l'acide nucléique de l'échantillon biologique ;

30 c) les réactifs nécessaires à la mise en œuvre d'une réaction d'amplification ;

d) les réactifs nécessaires à la détection et/ou au dosage de l'hybride formé entre ladite sonde et l'acide nucléique de l'échantillon biologique ou des acides nucléiques amplifiés formés.

Une telle trousse peut également contenir des contrôles positifs ou négatifs afin d'assurer la qualité des résultats obtenus. Elle peut également contenir les réactifs nécessaires à la purification des acides nucléiques à partir de l'échantillon biologique.

Le polynucléotide de l'Invention ou un de ses fragments, ainsi que les EST décrits précédemment ou leur fragments peuvent servir à la mise au point de modèles cellulaires ou animaux n'exprimant pas la protéine ASAP, en invalidant le gène ASAP par la méthode de Si RNA (ou RNAi pour RNA interference ; M. McManus and P. Sharp, Nature Reviews Genetics, 3, 737-747, 2002 ; V. Brondani, F. Kolb, E. Billy, M/S, 6-7, 665-667, 2002) à l'aide d'oligonucléotides dérivés de leurs séquences.

L'Invention a également pour objet un vecteur de clonage et/ou d'expression dans lequel est inséré le polynucléotide selon l'Invention.

Un tel vecteur peut contenir les éléments nécessaires à l'expression et éventuellement à la sécrétion de la protéine dans une cellule hôte.

Lesdits vecteurs comportent de préférence : un promoteur, des signaux d'initiation et de terminaison de la traduction, ainsi que des régions appropriées de régulation de la transcription. Ils doivent pouvoir être maintenus de façon stable dans la cellule et peuvent éventuellement comprendre des séquences codant pour des signaux particuliers spécifiant la sécrétion de la protéine traduite tels que par exemple un promoteur fort de nature ubiquitaire ou un promoteur sélectif d'un type de cellule et/ou de tissu particulier. Ces différentes séquences de contrôle sont choisies en fonction de l'hôte cellulaire utilisé.

Le polynucléotide selon l'Invention peut être inséré dans des vecteurs à réPLICATION autonome au sein de l'hôte choisi ou des vecteurs intégratifs de l'hôte choisi.

Parmi les systèmes à réPLICATION autonome, on utilise de préférence en fonction de la cellule hôte, des systèmes de type plasmidique ou viral. Les vecteurs viraux peuvent notamment être des adénovirus, des rétrovirus, des lentivirus, des poxvirus ou des virus herpétiques. L'homme du 5 métier connaît les technologies utilisables pour chacun de ces systèmes.

Lorsque l'on souhaite l'intégration de la séquence dans les chromosomes de la cellule hôte, on peut utiliser par exemple des systèmes de type plasmidique ou viral ; de tels virus sont, par exemple, les rétrovirus ou les virus associés aux adénovirus (Adeno-associated virus ou AAV).

10 Parmi les vecteurs non viraux, on préfère les polynucléotides nus tels que l'ADN ou l'ARN nu, les chromosomes artificiels de bactérie (BAC, bacterial artificial chromosome), les chromosomes artificiels de levure (YAC, yeast artificial chromosome) pour l'expression dans la levure, les chromosomes artificiels de souris (MAC, mouse artificial chromosome) pour l'expression 15 dans les cellules murines et de manière préférée les chromosomes artificiels d'homme (HAC, human artificial chromosome) pour l'expression dans les cellules humaines.

De tels vecteurs sont préparés selon les méthodes couramment utilisées par l'homme du métier, et les vecteurs recombinants en résultant 20 peuvent être introduits dans l'hôte approprié par des méthodes standards, telles que par exemple la lipofection, l'électroporation, le choc thermique, la transformation après perméabilisation chimique de la membrane, la fusion cellulaire.

L'Invention a aussi pour objet les cellules hôtes transformées, 25 notamment les cellules eucaryotes et procaryotes, dans lesquelles au moins un polynucléotide ou un fragment selon l'Invention ou au moins un vecteur selon l'Invention a été introduit.

Parmi les cellules utilisables aux sens de la présente Invention, on peut citer les cellules bactériennes, les cellules de levure, les 30 cellules animales, en particulier les cellules de mammifères ou encore les cellules végétales. On peut citer également les cellules d'insectes dans

lesquelles on peut utiliser des procédés mettant par exemple en oeuvre des baculovirus.

L'Invention a également pour objet les organismes transgéniques non-humains tels que les animaux ou les végétaux transgéniques, dont tout ou partie des cellules contient le polynucléotide selon l'Invention ou le vecteur selon l'Invention, sous une forme libre ou intégrée.

De préférence selon l'Invention, les organismes transgéniques non-humains sont ceux porteurs de cellules contenant un polynucléotide selon l'Invention, non fonctionnel ou porteur d'une mutation.

Selon l'Invention les animaux transgéniques sont de préférence des mammifères, excepté l'homme, plus préférentiellement les rongeurs, en particulier les souris ou les rats.

Les animaux transgéniques peuvent être obtenus par toute méthode classique connue de l'homme du métier, comme par exemple par recombinaison homologue sur cellules souches embryonnaires, transfert de ces cellules souches à des embryons, sélection des chimères affectées au niveau des lignées reproductrices, et croissance desdites chimères.

Les cellules hôtes transformées, les animaux ou les végétaux transgéniques selon l'Invention peuvent ainsi exprimer ou surexprimer le gène codant pour la protéine selon l'Invention ou leur gène homologue ou exprimer ledit gène dans lequel est introduite une mutation.

Les cellules de testicule ou de cerveau, les cellules hôtes transformées ou les organismes transgéniques selon l'Invention peuvent être utilisés pour la préparation de la protéine selon l'Invention.

La protéine selon l'Invention, particulièrement la protéine ASAP native, peut être purifiée selon les techniques connues de l'homme du métier. Ainsi, la protéine peut être purifiée à partir de lysats et extraits cellulaires, du surnageant du milieu de culture, par des méthodes utilisées individuellement ou en combinaison, telles que le fractionnement, les méthodes de chromatographie, particulièrement de chromatographie d'affinité, les techniques d'immunoaffinité à l'aide d'anticorps monoclonaux ou polyclonaux spécifiques, etc...

L'Invention a également pour objet une méthode de préparation de la protéine ASAP, se caractérisant en ce que l'on cultive des cellules exprimant la protéine ou des cellules transformées selon la présente Invention, notamment les cellules de mammifères ou les cellules d'organismes 5 transgéniques selon l'Invention, dans des conditions permettant l'expression de ladite protéine, et que l'on purifie ladite protéine.

Comme technique de purification, on peut citer par exemple la chromatographie d'affinité sur glutathione-sépharose (ou agarose) telle que décrite dans Sambrook J & Russell DW. (2000, Cold Spring Harbor 10 Laboratory Press).

L'Invention a également pour objet une protéine, caractérisée en ce qu'elle est susceptible d'être obtenue par l'une quelconque des méthodes de préparation ci-dessus décrites.

L'Invention a encore pour objet une méthode de criblage 15 d'une substance capable d'interagir *in vitro*, directement ou indirectement, avec le polynucléotide ou la protéine selon l'Invention caractérisée en ce que :

- dans une première étape on met en contact la substance à tester et le polynucléotide ou la protéine selon l'Invention et
- dans une deuxième étape on détecte par tout moyen approprié le complexe formé entre ladite substance et le polynucléotide ou la protéine selon l'Invention.

La présente invention a également pour objet une méthode de criblage d'une substance capable de moduler (activer ou inhiber) l'activité de la protéine ASAP, caractérisée en ce que :

- dans une première étape on met en contact des cellules d'un échantillon biologique exprimant la protéine ASAP avec une substance à tester,
- dans une deuxième étape on mesure par tout moyen approprié l'effet de ladite substance sur l'activité de ladite protéine ASAP, et
- dans une troisième étape on sélectionne des substances capables de moduler ladite activité.

Au sens de la présente invention, on entend par activité de la protéine ASAP, aussi bien l'expression de la protéine ASAP ou des transcrits (ARNm) correspondants, que l'activité biologique de ladite protéine ASAP, comme par exemple son effet sur l'organisation du fuseau mitotique ou 5 l'induction de mitoses aberrantes et abortives.

La détection du complexe formé entre ladite substance et le polynucléotide ou la protéine ou la mesure de l'effet de ladite substance sur l'activité de ladite protéine ASAP peuvent être réalisées par les techniques classiques d'analyse d'ARNm ou de protéines qui sont connues en 10 elles-mêmes ; à titre d'exemple non-limitatif, on peut citer les techniques suivantes : RT-PCR, Northern-blot, Western-blot, RIA, ELISA, immunoprécipitation, techniques d'analyse immunocytochimique ou immunohistochimique.

Avantageusement, ladite mesure est réalisée à l'aide des sondes, des amorces ou des anticorps, tels que définis ci-dessus.

De telles substances peuvent être des macromolécules biologiques comme par exemple un acide nucléique, un lipide, un sucre, une protéine, un peptide, un composé hybride protéine-lipide, protéine-sucre, peptide-lipide ou peptide-sucre, une protéine ou un peptide sur lequel on a ajouté des ramifications chimiques ou encore des molécules chimiques.

20 L'Invention a également pour objet le polynucléotide, la protéine, les anticorps, les vecteurs ou les cellules transformées selon l'Invention, utilisés comme médicaments.

Comme indiqué précédemment, la surexpression de la protéine selon l'Invention bloque la division cellulaire et par conséquent la 25 prolifération cellulaire. Cela en fait un excellent candidat pour une utilisation comme agent anti-mitotique, utilisable par exemple dans le traitement des pathologies cancéreuses.

Ainsi, l'Invention a également pour objet l'utilisation du polynucléotide, d'un vecteur ou de la protéine selon l'Invention, dans la préparation d'un médicament anti-mitotique.

De même comme il est également indiqué précédemment, la surexpression de la protéine selon l'Invention entraîne des perturbations dans

l'organisation du fuseau mitotique et induit des mitoses aberrantes et abortives (cellules pluriplurinucléées, fuseaux mono- ou multipolaires).

Ainsi, l'Invention a aussi pour objet l'utilisation d'un polynucléotide anti-sens ou d'un fragment anti-sens, d'un anticorps, d'un vecteur 5 contenant un oligonucléotide anti-sens selon l'Invention, capables d'inhiber l'expression du polynucléotide ou de la protéine selon l'Invention, dans la préparation d'un médicament destiné au traitement des pathologies liées aux perturbations dans l'organisation du fuseau mitotique et/ou à une induction des mitoses aberrantes et abortives (cellules pluriplurinucléées, fuseaux mono- ou 10 multipolaires) liées à la surexpression de la protéine selon l'Invention.

Outre les dispositions qui précèdent, l'Invention comprend encore d'autres dispositions qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en œuvre de l'Invention ainsi qu'aux dessins annexés, dans lesquels :

15 - La Figure 1 représente la localisation chromosomique et la structure du gène *asap* humain.

- La Figure 2 représente les signaux obtenus par Northern blots sur différents tissus humains après hybridation avec une sonde hASAP.

- La Figure 3 représente les résultats obtenus :

20 (A) par électrophorèse sur gel d'agarose des produits de RT-PCR obtenus avec des amorces correspondant au polynucléotide de souris, orthologue du polynucléotide SEQ ID NO: 15, à partir de différents tissus de souris.

(B) Après transfert du gel après électrophorèse sur une 25 membrane et hybridation avec une sonde mASAP interne.

- La Figure 4 représente la localisation cellulaire de la protéine hASAP couplée à la Green Fluorescent Protein (GFP) en 3' ou la Yellow Fluorescent Protein (YFP) en 5' ou à un tag MYC du côté N-terminal (colonne fusion).

30 Les noyaux sont colorés à l'iodure de propidium ou au Hoechst 33286. (4A : objectif 63x ; 4B, 4C, 4D : objectif 100X).

- La Figure 5 montre la co-localisation de la protéine ASAP

humaine avec l'alpha-tubuline. Figure 5 A : localisation cellulaire de l'alpha-tubuline, Figure 5 B : localisation de la protéine ASAP, Figure 5 C : superposition des 2 images montrant la colocalisation des 2 protéines.

Les exemples suivants sont illustratifs de l'Invention et ne la limitent aucunement.

EXEMPLE 1 : Construction de la séquence codante ASAP complète :

On amplifie la séquence complète de l'ADNc de la protéine ASAP à partir de 2 fragments chevauchants :

- un fragment A amplifié par PCR à partir du clone AI885274
10 avec les amores :

constFIS-1F (5'-ATGTCTGATGAAGTTTTAGCACC-3') (SEQ ID NO: 31) et

constFIS-2R (5'-AGGCCTCAAATGATGCTAATGC-3') (SEQ ID NO: 32) ;

15 - un fragment B amplifié à partir du clone AI671785 avec les amores :

constFIS-2F (5'-ATCATTGAGGCCTGGAAGGC-3') (SEQ ID NO: 33) et

et constFIS-1R (5'-AACACTTTGCGAACACAGTTC-3')
20 (SEQ ID NO: 34).

Puis, pour obtenir un produit PCR unique correspondant à la séquence complète de l'ADNc de la protéine ASAP, utilisable pour les expériences de fonction, 0,5 µl des produits de chacune des 2 réactions PCR (fragment A et B) sont hybridés ensemble à 25°C puis amplifiés avec les amores constFIS-1F et constFIS-2F. Ce produit PCR est sous-cloné dans le vecteur PCR4 suivant les recommandations du fabricant (Invitrogen) et vérifié par séquençage.

Les difficultés majeures rencontrées se sont situées dans la détermination *in silico* de la séquence codante complète ASAP et de sa reconstruction *in vitro*. En particulier, le choix des amores et des différentes PCR de la région 3' ont été délicats en raison de la richesse de la séquence en polyA.

EXEMPLE 2 : Analyse bio-informatique

La Figure 1 représente la localisation chromosomique et la structure du gène *asap* humain.

L'organisation complète du gène *asap* et sa localisation chromosomique ont été obtenues en comparant la séquence de l'ADNc obtenu à l'exemple 1, à la séquence du génome humain en utilisant les programmes du Wellcome Trust Sanger Institute (<http://www.ensembl.org/genome/central/> et plus précisément le programme de recherche BLAST (<http://genome.cse.ucsc.edu/>).

Le gène humain *asap* est constitué de 29750 nucléotides comprenant 14 exons dont 13 seulement sont traduits, le premier exon n'étant pas traduit. La taille des exons s'échelonne de 71 à 321 paires de bases. La séquence du gène est contenue dans le contig AC097467 (longueur 178204 paires de bases) entre les bases 115117 et 143828 (version v.7.29a3 NCBI/Ensembl du 12 juillet 2002, <http://www.ensembl.org>), et est par ailleurs localisée sur le chromosome 4q32.1 entre les marqueurs anonymes D4S1053 et D4S571 (région 161,25 Mégabases (Mb) à 161,28 Mb). La séquence du gène est physiquement contenue dans le clone BAC RP11-27G13.

Deux acides nucléiques correspondants à des fragments du polynucléotide isolé par les Inventeurs sont répertoriés dans la base de données GenBank sous les numéros d'accès AK024730 et AK024812, ainsi que les EST répertoriés sous les numéros d'accès BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289,

BB312835, AW347411, AA972439, BB263570, AU035125, BB277226,
BB274224, BB268445, AW024037, AA025609, BB274174, R96089,
BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277,
AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418,
5 BG591509, BF457670, AL897593, AL897592, BM926692, BM538559,
BI759567, AL601021, AL598780, AU222540, BG567619, AU166296,
BF889835, AU164011, AV656025, BF343454, AW262441, AW237952. Ces
séquences, obtenues dans le cadre d'un programme de séquençage en
masse de banques d'ADN complémentaires humains, sont incomplètes et
10 n'ont jamais été ni reconnues ni caractérisées.

La séquence protéique a été comparée aux séquences des banques de données en utilisant les programmes PSI-BLAST et PHI-BLAST du NCBI (<http://www.ncbi.nlm.nih.gov/Sitemap/>). Des motifs protéiques consensus ont été recherchés en utilisant les programmes DART du NCBI et 15 SMART d'ExPASy-Tools (<http://www.expasy.ch/tools/#similariw>), dont les paramètres permettent de détecter des motifs de faible homologie. La protéine ASAP présente une identité de séquence de 23% sur le tiers C-terminal avec une protéine associée aux microtubules (MAP 1A pour Microtubule-Associated-Protein 1A). Par ailleurs la recherche de motifs 20 conservés (DART on SMART) révèle des domaines de type caldesmon (Gusev, N.B., Biochemistry, 10 : 1112-1121, 2000) et ERM (Ezrin/radixin/moesin) (Louvet-Vallet, S., Biol. Cell, 274 : 305-316, 2000), qui sont des protéines également considérées comme des MAPs, avec des identités d'environ 20%. Elle présente également un domaine BRCT (Breast 25 cancer carboxy-terminal domain ; Bork, P., et coll., FASEB J., 11, 68-76 (1997)) entre les positions 65 et 303.

La protéine ASAP présente des domaines coiled-coil essentiellement regroupés dans sa partie C-terminale entre les acides aminés 297 et 327 d'une part et 477 et 628 d'autre part, indiquant soit que la protéine 30 s'oligomérise, soit qu'elle interagit avec d'autres protéines.

L'analyse informatique de la protéine à l'aide des programmes accessibles dans le site internet (<http://npsa-bil.ibcp.fr/cgi>-

bin/npsa_automat.pl?page=/NPSA/npsa_secons.html), révèle l'absence de feuillet β et une très grande richesse en hélices α , en particulier pour la région comprise entre les acides aminés 420-620, presque exclusivement formées d'hélices α .

5 Ces éléments permettent de penser que la protéine ASAP est une nouvelle MAP.

EXAMPLE 3 : Expression tissulaire

a) Analyse par Northern blot

Préparation des sondes radioactives :

10 Les ADN à radiomarquer sont isolés sur gel à bas point de fusion (LMP) selon la technique décrite dans Rouquier, S. et al., (Genomics, 17, 330-340, (1993)). Environ 100 ng d'ADN ainsi isolé, sont marqués par amorçage aléatoire (fragment de Klenow, Proméga) en présence de [α^{32} P dCTP] (Amersham) selon la technique décrite dans Feinberg, A.P. & 15 Vogelstein, B., (Anal. Biochem., 132, 6-13, (1983)). Ces sondes sont purifiées sur des colonnes de Sephadex G-50 selon la technique décrite dans Sambrook J & Russell DW. (2000, Cold Spring Harbor Laboratory Press). Les hybridations s'effectuent durant la nuit en présence de 2.10^6 Cpm/ml de sonde radioactive dénaturée.

20 a.1) Hybridation

Deux membranes Northern Blot de la société Clontech, (Human MTN Blot at Human MTN Blot II, Réf. 7760-1 et 7759-1) comportant des ARNm humains de différents tissus ont été hybridées avec l'ADNc hASAP complet marqué comme décrit ci-dessus. La membrane est hybridée en 25 présence de formamide à 42°C, en suivant le protocole Clontech. Un contrôle d'hybridation de la membrane est réalisé avec une sonde actine. La membrane est rincée 2 fois à haute stringence en 0,1X SSC/0,1% SDS à la température de 42°C, pendant 15 minutes. Les membranes sont alors analysées par autoradiographie ou au Phosphorimager.

30 Les tissus testés sont : la rate, le thymus, la prostate, le testicule, l'ovaire, l'intestin grêle, le colon, les leucocytes sanguins, le cœur, le

cerveau, le placenta, le poumon, le foie, le muscle squelettique, le rein et le pancréas.

a.2) Résultats

La Figure 2 illustre ces résultats.

5

Deux signaux sont détectés :

- un signal dans le testicule à environ 2,6 kb, ce qui correspond à la taille de l'ARNm ;

- un signal dans le cerveau mais à un haut poids moléculaire (9 kb) qui correspond soit à un prémessager, soit à une isoforme de haut poids moléculaire

10

b) Analyse par RT-PCR

15

Cette analyse a été effectuée sur des ARNs totaux de différents tissus de souris, à savoir le cerveau, le cœur, le colon, le foie, l'intestin grêle, le muscle squelettique, le pancréas, le poumon, le rein, la rate et le testicule.

b.1) Obtention de l'ADNc orthologue de souris

20

Les ARNs totaux de cellules de différents tissus de souris sont extraits avec le "mammalian total RNA kit" de la société Sigma. Les ARN sont rétro-transcrits avec le kit Superscript II de la société Invitrogen selon les conditions prescrites par le fournisseur et en utilisant des amores oligodT. Les produits obtenus sont vérifiés par électrophorèse sur gel d'agarose à 1%. 1 µl de chaque échantillon ainsi obtenu est à son tour amplifié par PCR (25 µl de milieu de réaction, 30 cycles (94°C pendant 15 secondes, 55°C pendant 30 secondes, 72°C pendant 30 secondes)) avec des amores spécifiques du gènes *asap* de souris (mFIS-1F, 5'-ACA ACG AAT AAC AGA GTG TCC-3' (SEQ ID NO: 35) et mFIS-2R, 5'-ACT CCT GAT AAA CAG CTG CC-3' (SEQ ID NO: 36)).

25

Les produits amplifiés obtenus sont analysés par électrophorèse sur gel d'agarose à 1%, colorés au bromure d'éthydium et leur taille comparée à un marqueur de taille déposé sur le gel en parallèle.

30

Après électrophorèse, les produits amplifiés obtenus sont transférés par capillarité sur membrane de nylon chargée, dans le tampon

NaCl 1,5M/NaOH 0,5M, selon la technique de Southern (transfert alcalin). La membrane est ensuite hybridée avec une sonde radiomarquée mASAP, (SEQ ID NO: 44), générée par amplification de la séquence contenue dans le clone de souris AW06131 sélectionné après comparaison de la séquence ASAP humaine dans les banques de données (GenBank) (http://expression.gnf.org/promoter/tissue/images/41739_s_at.png).

L'amplification a été réalisée par PCR (conditions telles que décrites ci-dessus dans lesquelles le volume de réaction est de 50 µl et le dCTP froid est à la concentration de 10 µM supplémenté avec 50 µCi d' α -P³²-dCTP à 3000Ci/mmol), en utilisant les amorces mFIS-1F (SEQ ID NO: 35) et mFis-2R (SEQ ID NO: 36). Les hybridations sont réalisées à 65°C (dans du tampon 6X SSC/0,5% SDS/5X Denhardt). La membrane est rincée à forte stringence (0,1X SSC/0,1% SDS), puis analysée par autoradiographie ou au PhosphorImager.

15 b.2) Résultats

La Figure 3 illustre ces résultats.

On constate qu'on obtient un signal majoritaire dans le testicule et le cerveau, nettement visible sur gel (Figure 3A).

Après transfert du gel et hybridation avec une sonde interne, 20 on constate que l'on détecte un signal très faible dans les autres tissus (Figure 3B).

Par conséquent l'ARNm codant pour la protéine mASAP est majoritairement exprimé dans le testicule et le cerveau. L'ADNc complet de souris, amplifié par RT-PCR à partir de l'ARN de testicule de souris, 25 correspond à la séquence SEQ ID NO : 45 et la protéine correspondante (mASAP) à la séquence SEQ ID NO : 46.

EXAMPLE 4 : Localisation cellulaire

a) Sous-clonage de l'ADNc hASAP dans un vecteur d'expression eucaryote

L'ADNc hASAP obtenu à l'exemple 1 est inséré dans trois 30 vecteurs d'expression :

1- dans pEAK10-EGFP en phase avec la Green Fluorescent Protein (GFP) fusionnée en C-terminal (vecteur 1) (pEAK10, vecteur de Edge

Biosystems (distribué par Q.BIOgene, Illkirch en France) dans lequel a été introduit la protéine EGFP (enhanced Green Fluorescent Protein) suivant la référence Gaillard, I., et al., Eur. J. Neurosci., 15, 409-418, 2002) ;

2- dans pEYFP-C1 en phase avec la Yellow Fluorescent Protein (YFP) fusionnée du côté N-terminal (vecteur 2) (distribué par BD Biosciences Clontech)) ;

3- dans GLOMYC3-1 comportant un tag MYC du côté N-terminal (vecteur 3), vecteur dérivé du vecteur pcDNA3.1 (Invitrogen), dans lequel ont été insérées une région 5' non-traduite (5' UTR) et un tag MYC aux sites *HindIII-BamHI*, et la région 3'UTR de la globine (fragment *SpeI-XbaI* dans le site *XbaI*).

L'ADNc hASAP est amplifié à partir de son vecteur de clonage initial (pCR4-TOPO) par PCR en utilisant la polymérase haute-fidélité pfu Turbo, à l'aide d'amorces amplifiant l'ADNc entre la méthionine de départ et le dernier acide aminé. Les produits amplifiés obtenus sont sous-clonés dans les 3 vecteurs.

- Clonage dans PEAK-GFP. Préparation de l'insert d'ADN par PCR [94°C 2 min ; (94°C 15 sec.; 58°C 30 sec.; 72°C 1min 30 sec.) 30 cycles ; 72°C 3 min.], à l'aide des amorces hFIS-Exp1F (5'-GCCACCATGTCTGATGAAGTTTTAGCAC-3) (SEQ ID NO: 37) et hFIS-Exp1R (5'-GAAACACTTTGCGAACACAGTTC-3') (SEQ ID NO: 38).

Le vecteur est coupé par *EcoRV* et déphosphorylé : 10 ng de vecteur sont utilisés pour la ligation avec l'insert d'ADN. Le produit PCR est phosphorylé puis purifié sur high PURE PCR kit (Roche) : 100 ng d'insert sont utilisés pour la ligation [12h à 16°C dans 10 µl final (ligase Biolabs), suivant les conditions standards (Sambrook and Russell)].

- Clonage dans Glomyc : Préparation de l'insert d'ADN par PCR [94°C 2 min ; (94°C 15 sec.; 60°C 30 sec.; 72°C 1min 30 sec.) 30 cycles ; 72°C 3 min.], à l'aide des amorces : Glomyc-FIS1F : (5'-TAATGTCTGATGAAGTTTTAGCACC-3') (SEQ ID NO: 39) et

Glomyc-FIS1R : (5'-TCAAAACACTTTGCGAACACAGTTC-3') (SEQ ID NO: 40).

Conditions de clonage identiques à celles décrites pour le clonage dans PEAK-GFP.

5 - Clonage dans YFP : Préparation de l'insert d'ADN : mêmes conditions que pour Glomyc, à l'aide des amorces :

YFP-FIS1F (5'-AATGTCTGATGAAGTTTTAGCACC-3') (SEQ ID NO: 41) et Glomyc-FIS1R (SEQ ID NO: 40) (cf ci-dessus).

10 Conditions de clonage identiques à celles décrites pour le clonage dans PEAK-GFP, le vecteur ayant préalablement été coupé par *Sma*1.

Les recombinants sont analysés par PCR en utilisant une amorce du vecteur et une amorce interne.

PEAK-GFP : annealing à 58°C, extension 45 sec. à 72°C. et conditions 15 standards pour le reste. Amorces : constFIS-2F (SEQ ID NO: 33) et GFP-1R (5'-TCAGCTTGCCGTAGGTGGC-3') (SEQ ID NO: 42).

YFP : annealing 55°C pendant 1 min. ; Amorces : YFP-2F (5'-ATGGTCCTGCTGGAGTTCG-3') (SEQ ID NO: 43) et hFIS-Exp1R (SEQ ID NO: 38) .

20 Glomyc : annealing 44°C, extension 45 sec. à 72°C. Amorces : constFIS-2F (SEQ ID NO: 33) et SP6. Les recombinants sont séquencés par séquençage automatique à façon à partir des produits PCR (Genome Express, Meylan).

b) Sous-clonage de l'ADNc hASAP dans un vecteur d'expression procaryote

25 En utilisant une stratégie similaire à celle utilisée au paragraphe a) ci-dessus, l'ADNc hASAP a été cloné dans le vecteur pGEX-4T2 (AMERSHAM), de façon à produire une protéine de fusion avec la GST, purifiable selon les protocoles standards.

c) Sous-clonage de l'ADNc mASAP dans un vecteur d'expression procaryote ou eucaryote

30 En utilisant une stratégie similaire à celle utilisée au paragraphe a) ci-dessus, l'ADNc mASAP a été cloné dans les vecteurs suivants :

- pGEX-4T2, (AMERSHAM), de façon à produire une protéine de fusion avec la GST, purifiable selon les protocoles standards.

- pEYFP-C1 de façon à produire une protéine de fusion (fusion N-terminale) avec la Yellow Fluorescent Protein (YFP) détectable par immunofluorescence directe.

5 d) Transfection, immunofluorescence et microscopie

d.1) matériels et méthodes

Les vecteurs obtenus sont transfectés selon la technique au phosphate de calcium ou de façon plus routinière en utilisant le procédé 10 jetPEI (GDSP10101, Qbiogene) suivant les recommandations du fabricant, dans les lignées cellulaires suivantes :

- PEAK (ref. 37937, Edge Biosystems (distribué par Q.BIOgene, Illkirch en France), uniquement pour les constructions ASAP humaines,

15 - HEK-293 (ATCC (American Tissue Culture Collection) référence CRL-1573 ; p53 +/- non synchronisable), pour les constructions ASAP humaines et murines,

- NIH3T3 non-transformées (constructions ASAP murines), et
- U-2 OS (ATCC HTB-96 ; p53 +/-, synchronisable)

20 Pour les vecteurs 1) et 2) (constructions ASAP humaines et murines), les localisations se font directement par détection de la fluorescence de la GFP ou de l'YFP à 24h, 48h et 72h après fixation des cellules au paraformaldéhyde et coloration des noyaux soit au propioiodure de propidium, soit au Hoechst 33286.

25 Pour le vecteur 3) la détection du tag MYC est réalisée à l'aide d'un anticorps primaire anti-MYC distribué par TEBU (9 E10, cat.#SC-40, Santa Cruz Biotechnology, CA) et d'un anticorps secondaire de chèvre anti-IgG de souris marqué au fluorochrome Alexa-594 (Molecular Probes, ref. A-11032, distribué en France par Interchim, Montluçon), après la fixation des 30 cellules et leur perméabilisation au Triton X100 0,1%. Les lames sont analysées, et les images collectées sur un microscope Zeiss Axiophot.

d.2) Résultats : Localisation cellulaire et colocalisation de la protéine ASAP avec l'alpha-tubuline
- localisation cellulaire

La Figure 4 illustre la localisation cellulaire de la protéine 5 hASAP surexprimée dans la lignée HEK-293 (IP = Iodure de propidium).

L'observation au microscope à fluorescence des lames correspondant aux différentes transfections par les vecteurs 1), 2) et 3) montrent les mêmes types de profil : la localisation des protéines hASAP et mASAP est cytoplasmique et son profil fibreux rappelle celui des filaments de 10 tubuline.

Par ailleurs, il semble que les cellules transfectées présentent des défauts de division car les noyaux sont toujours plus gros que dans les cellules non transfectées (Figure 4A et 4B). De plus, certaines des cellules transfectées semblent pluri nucléées (Figure 4B). Ceci suggère une division 15 anormale des cellules transfectées.

Enfin, la mitose des cellules transfectées semble anormale, tant au niveau de l'organisation des chromosomes, que du profil de localisation des protéines hASAP et mASAP au niveau du fuseau mitotique. Le profil de localisation des protéines hASAP et mASAP, en étoile, est caractéristique 20 de la nucléation des microtubules en aster autour du centrosome (Figures 4C et 4D).

Un profil similaire de localisation de la protéine ASAP est détecté dans la lignée U-2 OS (p53 +/-) surexprimant hASAP et dans la lignée NIH 3T3 non-transformée surexprimant mASAP ; une accumulation de 25 cellules monopolaires en mitose est observée.

En outre, en synchronisant les cellules U-2 OS et en récupérant les extraits cellulaires à différents moments du cycle, il a été vérifié que la protéine ASAP était bien présente dans toutes les phases du cycle cellulaire (interphase, S, G2/M).

30 - colocalisation de la protéine ASAP avec l'alpha-tubuline

La Figure 5 illustre la co-localisation de la protéine ASAP humaine avec l'alpha-tubuline ; de même la protéine ASAP murine co-localise

avec l'alpha-tubuline.

La Figure 5 A illustre la localisation cellulaire de l'alpha-tubuline détectée par immunofluorescence à l'aide d'un anticorps anti-tubuline (Alexa-594, Molecular Probe).

5 La Figure 5 B illustre la localisation de la protéine ASAP marquée à la YFP (yellow fluorescent protein).

La Figure 5 C représente la superposition des 2 images démontrant la colocalisation des 2 protéines.

EXAMPLE 5 : Production d'anticorps polyclonaux anti-hASAP et mASAP

10 a) Production d'anticorps

Les constructions suivantes de la protéine ASAP ont été clonées dans le vecteur d'expression procaryote pGEX 4T-2 (AMERSHAM) comme décrit à l'exemple 4 :

- 15 - protéine ASAP humaine entière (SEQ ID NO : 1)
- protéine humaine délétée de sa partie C-terminale contenant
le domaine MAP potentiel (résidus 1 à 421, SEQ ID NO : 47)
- protéine murine entière (SEQ ID NO : 46).

20 Les protéines ont été exprimées dans *E. coli* et purifiées selon les protocoles standards. Des lapins ont ensuite été immunisés avec les protéines ASAP purifiées selon un protocole standard, et les sérums immuns ont été récoltés.

b) Analyse de la réactivité des sérums polyclonaux vis-à-vis de la protéine ASAP endogène.

25 Les sérums polyclonaux monospécifiques dirigés contre la protéine hASAP entière ou délétée de sa partie C-terminale contenant le domaine MAP potentiel, ont été testées en Western-blot et en immunofluorescence, sur des cellules HEK-293 et U-2 OS, selon des protocoles standards.

30 En Western-blot, le sérum polyclonal monospécifique dirigé contre la protéine hASAP entière détecte une protéine d'un poids moléculaire apparent d'environ 110 kDa correspondant à la protéine ASAP endogène, aussi bien dans les cellules HEK-293 que les cellules U-2 OS. Dans ces conditions, un anticorps anti-FLAG, détecte une protéine de poids moléculaire

équivalent, dans des cellules contrôle HEK-293 ou U-2 OS, transfectées par un vecteur d'expression de la protéine hASAP fusionnée avec une étiquette FLAG.

En immunofluorescence, le sérum polyclonal monospécifique dirigé contre la protéine hASAP entière marque les microtubules des cellules HEK-293 en interphase, les asters des cellules en mitose et les microtubules du corps résiduel en fin de télophase.

Le sérum polyclonal monospécifique dirigé contre la protéine hASAP délétée de sa partie C-terminale contenant le domaine MAP potentiel présente le même profil en immunofluorescence et détecte une protéine d'environ 110 kDa, en Western blot.

Le sérum polyclonal monospécifique dirigé contre la protéine mASAP est utilisé pour détecter quels sont les types cellulaires exprimant ASAP et à quel(s) stade(s) du cycle cellulaire elle est exprimée, par immunofluorescence sur des coupes testiculaires de souris.

EXAMPLE 6 : Analyse fonctionnelle de la protéine hASAP à l'aide de mutants délétés de la partie N-terminale contenant le domaine BRCT ou de la région C-terminale contenant le domaine MAP potentiel.

Des fragments d'ADNc codant pour une protéine hASAP délétée de la partie N-terminale contenant le domaine BRCT (Ndel1 : résidus 304 -647 (SEQ ID NO : 48) ; Ndel2 : résidus 411-647 (SEQ ID NO : 49) ; Ndel3 : résidus 478-647 (SEQ ID NO : 50)) ou de la partie C-terminale contenant le domaine MAP (Cdel1 : résidus 1 à 477 (SEQ ID NO : 51) ; Cdel2 : résidus 1 à 418 (SEQ ID NO : 52) ; Cdel3 : résidus 1 à 303 (SEQ ID NO : 53)) ont été amplifiés par PCR à l'aide d'amorces appropriées puis clonés dans les vecteurs d'expression pEAK10-EGFP (fusion C-terminale avec la GFP) et pEYFP-C1 (fusion N-terminale avec la YFP) selon un protocole similaire à celui décrit à l'exemple 4.

Les différentes constructions ont été transfectées dans les lignées HEK-293 et U-2 OS puis la localisation cellulaire des différents mutants de la protéine hASAP a été analysée comme décrit à l'exemple 4.

On constate que pour les mêmes délétions, un profil similaire est obtenu avec la construction comportant la YFP en N-terminal ou la GFP en C-terminal.

Par comparaison avec la protéine hASAP entière, les 3 constructions délétées de la partie C-terminale ne colocalisent plus en interphase avec la tubuline et ne présentent plus un aspect fibreux ; ces résultats indiquent que la délétion intéresse un domaine MAP. En outre, aucune cellule monopolaire bloquée en mitose n'est observée dans les cellules surexprimant les mutants délétés de la partie C-terminale contenant le domaine MAP.

Par comparaison avec la protéine hASAP entière, les 3 constructions délétées de la partie N-terminale contenant le domaine BRCT, présentent une localisation nucléaire sous forme de foyers mais il reste dans le cytoplasme quelques fibres co-localisant avec la tubuline.

L'analyse fonctionnelle de la protéine hASAP est complétée par des expériences d'inactivation de l'expression du gène par des ARN interférents (ARNi).

REVENDICATIONS

1°) Protéine isolée, dénommée ASAP, caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par :

5 a) une protéine répondant à la séquence représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO: 1, correspondant à la protéine ASAP humaine, et

10 b) une protéine présentant, sur la totalité de sa séquence, au moins 80% d'identité ou au moins 90% de similarité, de préférence au moins 90 % d'identité ou moins 95% de similarité, avec la protéine de séquence SEQ ID NO: 1.

2°) Protéine selon la revendication 1, caractérisée en ce qu'elle présente la séquence SEQ ID NO: 46 correspondant à la protéine ASAP murine.

15 3°) Peptide, caractérisé en ce qu'il est constitué d'un fragment d'au moins 10 acides aminés consécutifs d'une protéine selon la revendication 1 ou la revendication 2.

20 4°) Peptide selon la revendication 3, caractérisé en ce qu'il est sélectionné parmi les séquences représentées dans la liste de séquence en annexe sous les numéros SEQ ID NO: 2 à SEQ ID NO: 14 et SEQ ID NO: 47 à 53.

5°) Protéine variante de la séquence SEQ ID NO: 1 ou de la séquence SEQ ID NO: 46, caractérisée en ce qu'elle présente une mutation entraînant un dysfonctionnement de la protéine.

25 6°) Protéine ou peptide selon l'une quelconque des revendications 1, 3, 4 ou 5, caractérisée en ce qu'il s'agit d'une protéine ou d'un peptide d'origine humaine.

7°) Polynucléotide isolé, répondant à une séquence sélectionnée dans le groupe constitué par :

30 - les séquences codant pour une protéine ou pour un peptide selon l'une quelconque des revendications 1 à 6 ;

- les séquences représentées sous les numéros SEQ ID NO: 15 et SEQ ID NO: 45 dans la liste de séquences en annexe, correspondant respectivement aux ADNc ASAP humain et murin,
- le fragment d'ADN génomique de 29800 nucléotides 5 répondant à la séquence représentée dans la liste des séquences en annexe sous le numéro SEQ ID NO: 16, correspondant au gène *asap* humain ;
 - un fragment d'au moins 15 nucléotides consécutifs de l'une quelconque des séquences précédentes, à l'exclusion des séquences répertoriées sous les numéros d'accès AK024730 et AK024812 et des EST 10 répertoriés sous numéros d'accès BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, 15 AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, 20 BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, 25 BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952 dans la base de données GenBank ;
 - un fragment de l'un quelconque des séquences précédentes, sélectionné parmi les séquences représentées dans la liste des séquences en annexe sous les numéros SEQ ID NO: 16 à SEQ ID NO: 30 ;

- une séquence présentant un pourcentage d'identité d'au moins 80 %, de préférence 90 %, après alignement optimal avec l'une des séquences ou l'un des fragments précédents ;
 - les séquences complémentaires des précédentes, sens ou 5 anti-sens.

8°) Polynucléotide ou fragment selon la revendication 7, caractérisé en ce qu'il s'agit d'un polynucléotide variant de la séquence SEQ ID NO: 15 ou 45.

9°) Polynucléotide ou fragment selon la revendication 8, 10 caractérisé en ce qu'il s'agit d'un polynucléotide porteur d'au moins une mutation conduisant à une modification de la séquence en acides aminés de la protéine correspondant à la séquence SEQ ID NO: 1 ou SEQ ID N : 46.

10°) Utilisation d'un polynucléotide ou d'un fragment selon l'une quelconque des revendications 7 à 9 ou de l'une des séquences répertoriées sous les numéros d'accès AK024730 et AK024812 et d'un des EST répertoriés sous les numéros d'accès BU198882, BM693711, AW372449, 15 BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, 20 BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, 25 BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, 30 BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, 35 AL601021, AL598780, AU222540, BG567619, AU166296, BF889835,

AU164011, AV656025, BF343454, AW262441, AW237952 dans la base de données GenBank ou de leurs fragments, comme sonde pour détecter, identifier ou doser des polynucléotides correspondants au polynucléotide selon l'une quelconque des revendications 7 à 9, particulièrement dans d'autres organismes.

11°) Utilisation selon la revendication 10, caractérisée en ce que la sonde est sélectionnée dans le groupe constitué par les séquences SEQ ID NO: 15, 45 ou SEQ ID NO: 17 à SEQ ID NO: 44.

12°) Amorce pour l'amplification des polynucléotides (ARN ou ADN génomique) correspondants au polynucléotide selon l'une quelconque des revendications 7 à 9, particulièrement dans d'autres organismes caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par les séquences SEQ ID NO: 31 à 43.

13°) Polynucléotide susceptible d'être obtenu par amplification à l'aide des amorces selon la revendication 12.

14°) Méthode de détermination du profil de transcription du gène correspondant au polynucléotide selon l'une quelconque des revendications 7 à 9 ou 13, ou d'une altération dudit profil, dans un échantillon biologique, comprenant une première étape d'obtention par tout moyen approprié des ARN totaux à partir de l'échantillon biologique, une deuxième étape de mise en contact desdits ARN avec une sonde telle que définie aux revendications 10 ou 11 préalablement marquée dans des conditions classiques d'hybridation entre les ARN et la sonde et une troisième étape de révélation par tout moyen approprié des hybrides formés.

15°) Méthode selon la revendication 14, dans laquelle la deuxième étape est une étape de transcription inverse et/ou d'amplification des transcrits, réalisée à l'aide d'une paire d'amorces selon la revendication 12 et la troisième étape est une étape de révélation par tout moyen approprié des acides nucléiques amplifiés.

16°) Méthode selon l'une quelconque des revendications 14 ou 15, caractérisée en ce qu'elle comporte en outre une étape d'évaluation du

taux de transcription du gène par comparaison à un témoin préalablement choisi.

17°) Méthode de mise en évidence dans d'autres espèces du gène correspondant au polynucléotide selon l'une quelconque des revendications 7 à 9 ou 13, ou de mise en évidence des variants alléliques dudit gène ou de mise en évidence d'une altération fonctionnelle de ce gène, dans un échantillon biologique, comprenant une première étape d'obtention par tout moyen approprié de l'ADN à partir de l'échantillon biologique, une deuxième étape de mise en contact desdits ADN avec une sonde telle que définie aux revendications 10 ou 11, préalablement marquée, dans des conditions classiques d'hybridation entre les ADN et la sonde et une troisième étape de révélation par tout moyen approprié des hybrides formés.

18°) Méthode selon la revendication 17, dans laquelle la deuxième étape est une étape d'amplification réalisée à l'aide d'amorces selon la revendications 12 et la troisième étape une étape de révélation par tout moyen approprié des acides nucléiques amplifiés formés.

19°) Méthode selon l'une quelconque des revendications 17 ou 18, caractérisée en ce qu'elle comporte en outre une étape d'isolement et de séquençage des acides nucléiques mis en évidence.

20 20°) Trousse de réactifs pour la mise en œuvre des méthodes selon l'une quelconque des revendications 14 à 19 comprenant :

- au moins une sonde telle que définie aux revendications 10 ou 11 et/ou des amorces selon la revendication 12 ;
- les réactifs nécessaires à la mise en œuvre d'une réaction classique d'hybridation entre ladite sonde et/ou lesdites amorces et l'acide nucléique de l'échantillon biologique ;
- les réactifs nécessaires à la mise en œuvre d'une réaction d'amplification ;
- les réactifs nécessaires à la détection et/ou le dosage de l'hybride formé entre ladite sonde et l'acide nucléique de l'échantillon biologique ou des acides nucléiques amplifiés formés.

21°) Vecteur de clonage et/ou d'expression dans lequel est inséré un polynucléotide ou un fragment selon l'une quelconque des revendications 7 à 9 ou 13.

5 22°) Cellules hôtes transformées, dans lesquelles au moins un polynucléotide ou un fragment selon l'une quelconque des revendications 7 à 9 ou 13 ou au moins un vecteur selon la revendication 21 a été introduit.

10 23°) Organismes transgéniques non-humains, dont tout ou partie des cellules contient au moins un polynucléotide ou d'un fragment selon l'une quelconque des revendications 7 à 9 ou 123 ou au moins un vecteur selon la revendication 20, sous une forme libre ou intégrée.

24°) Organismes transgéniques non-humains selon la revendication 23 caractérisés en ce qu'ils sont porteurs de cellules contenant un polynucléotide selon l'une quelconque des revendications 7 à 9 ou 13 non fonctionnel ou porteur d'une mutation.

15 25°) Utilisation d'une cellule transformée selon la revendication 22 ou d'un organisme transgénique non-humain selon l'une quelconque des revendications 23 ou 24, pour la production d'une protéine ou d'un peptide selon l'une quelconque des revendications 1 à 6.

20 26°) Méthode de préparation d'une protéine ou d'un peptide selon l'une quelconque des revendications 1 à 6, caractérisée en ce que l'on cultive des cellules exprimant la protéine ou des cellules transformées selon la revendication 22 ou un organisme transgénique selon l'une quelconque des revendications 23 ou 24, dans des conditions permettant l'expression de ladite protéine, et que l'on purifie ladite protéine.

25 27°) Protéine, caractérisée en ce qu'elle est susceptible d'être obtenue par la méthode de préparation selon la revendication 26.

28°) Anticorps mono- ou polyclonal caractérisé en ce qu'il est capable de reconnaître spécifiquement une protéine ou un peptide selon l'une quelconque des revendications 1 à 6 ou 27.

30 29°) Anticorps selon la revendication 28, caractérisé en ce qu'il reconnaît, parmi les MAPs, uniquement et spécifiquement la protéine ASAP de séquence SEQ ID NO: 1 ou SEQ ID NO: 46.

30°) Utilisation d'un anticorps selon l'une quelconque des revendications 28 ou 29 pour la détection et/ou la purification d'une protéine selon l'une quelconque des revendications 1 à 6 ou 27.

5 31°) Méthode de détection d'une protéine selon l'une quelconque des revendications 1 à 6 ou 27, dans les cellules d'un échantillon biologique, comprenant

- une première étape de traitement convenable des cellules par tout moyen approprié permettant de rendre accessible le milieu intracellulaire,

10 - une seconde étape de mise en contact dudit milieu intracellulaire ainsi obtenu avec un anticorps selon l'une quelconque des revendications 28 ou 29 et

- une troisième étape de mise en évidence par tout moyen approprié du complexe protéine ASAP-anticorps formé.

15 32°) Utilisation d'un anticorps selon l'une quelconque des revendications 28 ou 29 pour la détection et/ou la sélection de cellules présentant des perturbations dans l'organisation du fuseau mitotique et/ou une induction des mitoses aberrantes et abortives liées à la surexpression de la protéine selon l'une quelconque des revendications 1 à 6 ou 27.

20 33°) Méthode d'évaluation *in vitro* de la capacité de prolifération ou d'agressivité de cellules cancéreuses comprenant

- une première étape de traitement convenable des cellules par tout moyen approprié permettant de rendre accessible le milieu intracellulaire,

25 - une seconde étape de mise en contact dudit milieu intracellulaire ainsi obtenu avec un anticorps selon l'une quelconque des revendications 28 ou 29,

- une troisième étape de mise en évidence et de mesure par tout moyen approprié du complexe protéine ASAP-anticorps formé et

30 - une quatrième étape d'évaluation du taux de transcription du gène par comparaison du taux de complexes protéine ASAP-anticorps formés à celui d'un échantillon biologique témoin préalablement choisi.

34°) Trousse permettant de mettre en oeuvre la méthode selon l'une quelconque des revendications 31 ou 33 comprenant :

- au moins un anticorps monoclonal ou polyclonal selon l'une quelconque des revendications 28 ou 29 ;
- 5 - les réactifs permettant la détection du complexe protéine ASAP-anticorps produit lors de la réaction immunologique.

35°) Méthode de criblage d'une substance capable d'interagir *in vitro*, directement ou indirectement, avec le polynucléotide selon l'une quelconque des revendications 7 à 9 ou 13 ou la protéine ou un peptide selon 10 l'une quelconque des revendications 1 à 6 ou 27, caractérisée en ce que :

- dans une première étape on met en contact la substance à tester et le polynucléotide ou la protéine et
- dans une deuxième étape on détecte par tout moyen approprié le complexe formé entre ladite substance et le polynucléotide ou la 15 protéine.

36°) Méthode de criblage d'une substance capable de moduler l'activité de la protéine selon l'une quelconque des revendications 1 à 6 ou 27, caractérisée en ce que :

- dans une première étape on met en contact des cellules d'un échantillon biologique exprimant la protéine selon l'une quelconque des revendications 1 à 6 ou 27, avec une substance à tester,
- dans une deuxième étape on mesure par tout moyen approprié de l'effet de ladite substance sur l'activité de ladite protéine, et
- dans une troisième étape on sélectionne des substances 25 capables de moduler ladite activité.

37°) Protéine selon l'une quelconque des revendications 1 à 6 ou 27 ou polynucléotide ou fragment selon l'une quelconque des revendications 7 à 9 ou 13 ou anticorps selon l'une quelconque des revendications 28 ou 29 ou vecteur selon la revendication 21 ou cellule transformée selon la 30 revendication 22, utilisé comme médicaments.

38°) Utilisation d'une protéine selon l'une quelconque des revendications 1 à 6 ou 27 ou d'un polynucléotide ou d'un fragment selon

l'une quelconque des revendications 7 à 9 ou 13 ou d'un anticorps selon l'une quelconque des revendications 28 ou 29 ou d'un vecteur selon la revendication 21 ou d'une cellule transformée selon la revendication 22, dans la préparation d'un médicament anti-mitotique.

- 5 39°) Utilisation d'un polynucléotide ou d'un fragment anti-sens selon l'une quelconque des revendications 7 à 9 ou 13 ou d'un anticorps selon l'une quelconque des revendications 28 ou 29 ou d'un vecteur contenant un oligonucléotide anti-sens selon la revendication 21 et capable d'inhiber l'expression du polynucléotide selon l'une quelconque des revendications 7 à
10 9 ou de la protéine selon l'une quelconque des revendications 1, 2 ,5, 6 dans la préparation d'un médicament destiné au traitement des pathologies liées aux perturbations dans l'organisation du fuseau mitotique et/ou à une induction des mitoses aberrantes et abortives, liées à la surexpression de la protéine selon l'une quelconque des revendications 1 à 6 ou 27.

1/5

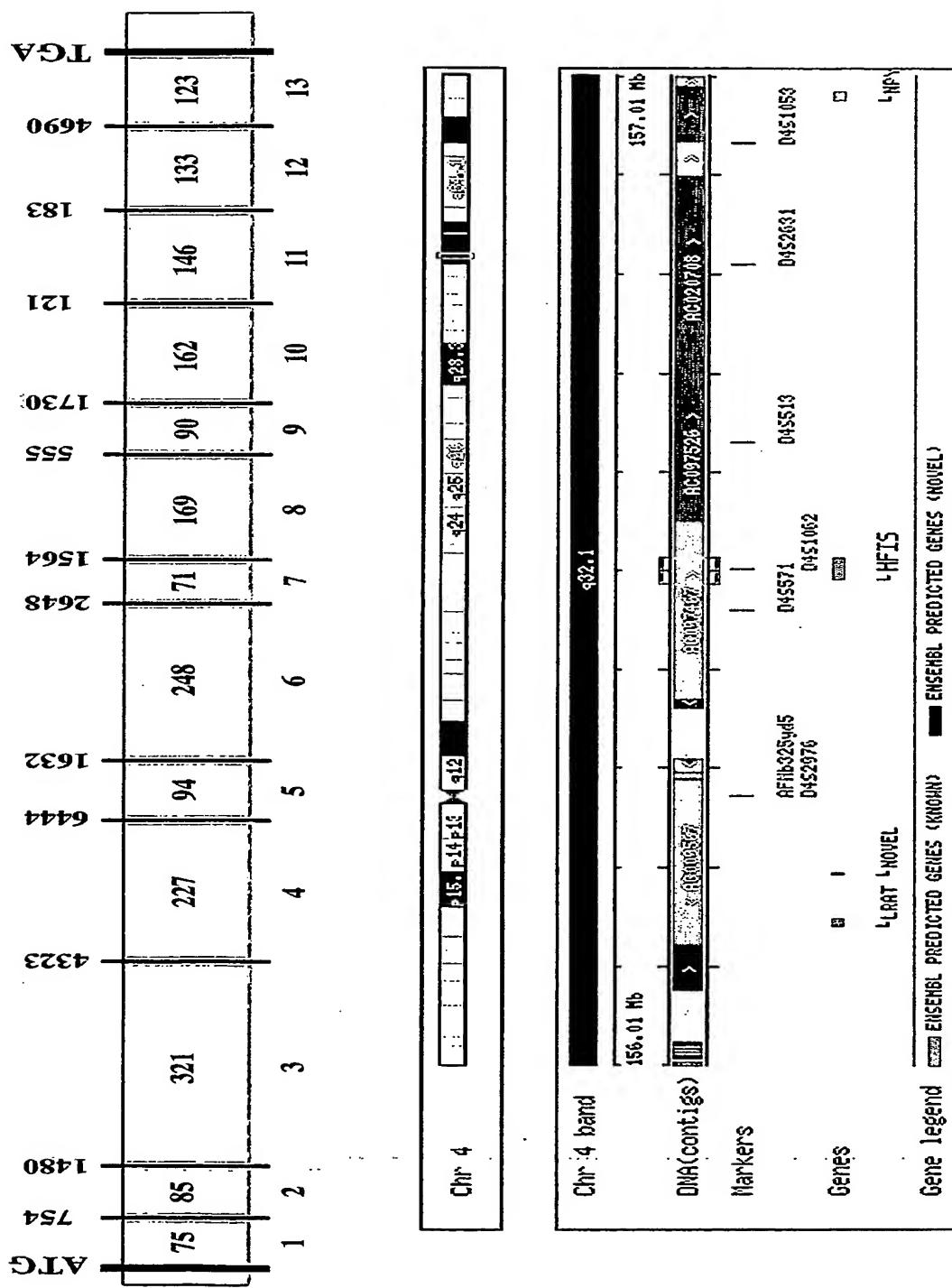


Figure 1

JOCO Record PCT/US 24 JUN 2005

THIS PAGE BLANK (USPTO)

2/5

Rate	
Thymus	
Prostate	
Testicule	
Ovaire	
Intestin grêle	
Colon	
Leucocyte du sang périphérique	
Coeur	
Cerveau	
Placenta	
Poumon	
Liver	
Muscle squelettique	
Rein	
Pancreas	

9 kb -

2,6 kb

Figure 2

JCO9 Rec'd PCT/PTO 24 JUN 2005

THIS PAGE BLANK (USPTO)

3/5

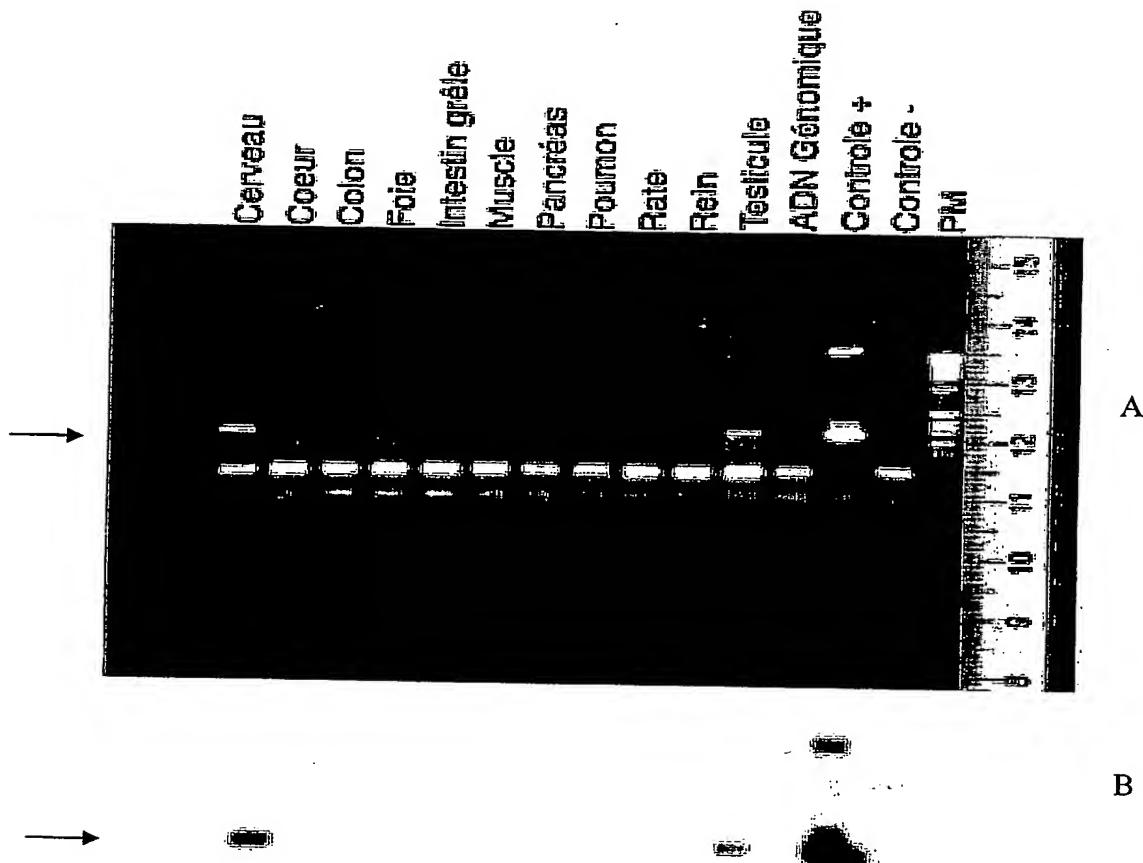
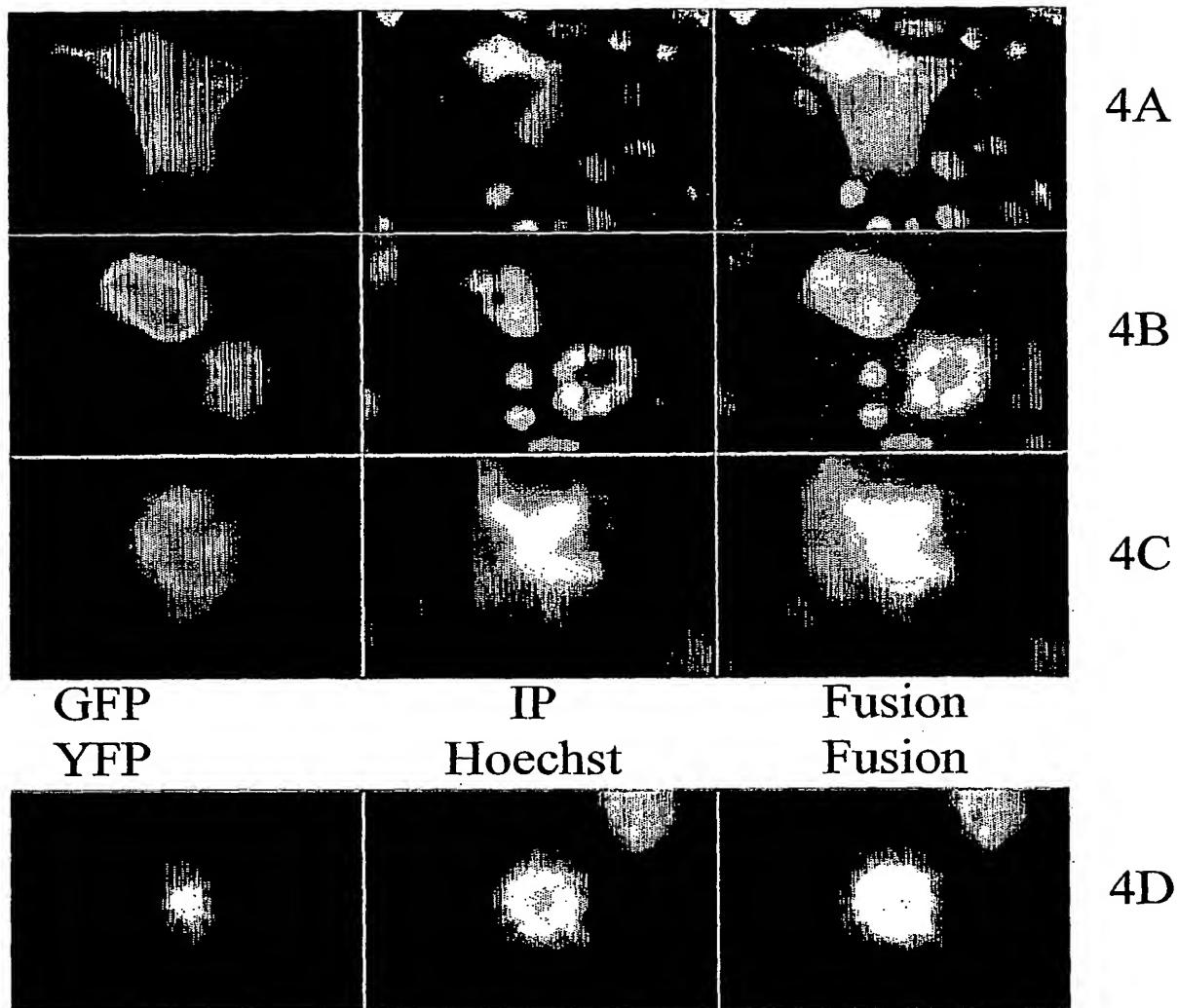


Figure 3

JC09 Rec'd PCT/PTO 24 JUN 2005

THIS PAGE BLANK (USPTO)

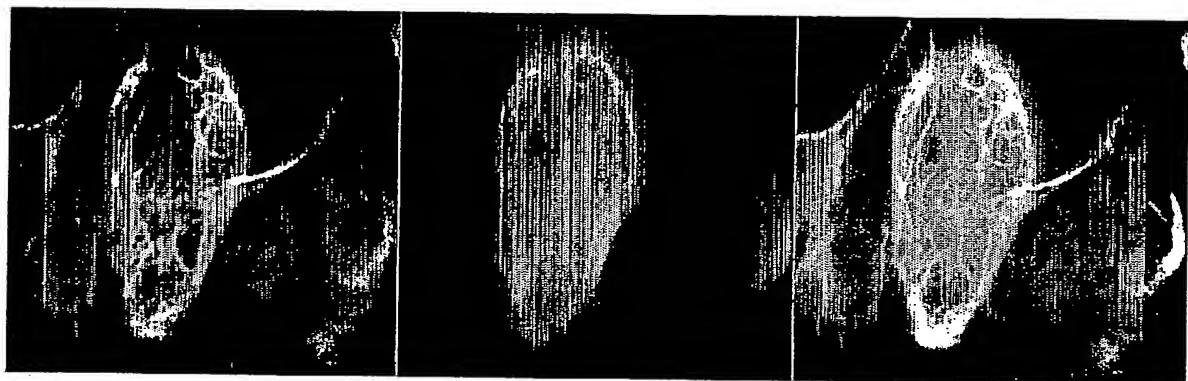
4/5

**Figure 4**

JC09 Rec'd PCT/PTO 24 JUN 2005

THIS PAGE BLANK (USPC)

5/5



A

B

C

Figure 5

THIS PAGE BLANK (over)

WO 2004/058815

s644PCT88.ST25
SEQUENCE LISTING

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

GIORGI, Dominique

SAFFIN, Jean-Michel

ROUQUIER, Sylvie

<120> Nouvelle protéine associée aux centrosomes et ses applications

<130> s644PCT88

<160> 53

<170> PatentIn version 3.1

<210> 1

<211> 647

<212> PRT

<213> Homo sapiens

<400> 1

Met Ser Asp Glu Val Phe Ser Thr Thr Leu Ala Tyr Thr Lys Ser Pro
1 5 10 15Lys Val Thr Lys Arg Thr Phe Gln Asp Glu Leu Ile Arg Ala Ile
20 25 30Thr Ala Arg Ser Ala Arg Gln Arg Ser Ser Glu Tyr Ser Asp Asp Phe
35 40 45Asp Ser Asp Glu Ile Val Ser Leu Gly Asp Phe Ser Asp Thr Ser Ala
50 55 60Asp Glu Asn Ser Val Asn Lys Lys Met Asn Asp Phe His Ile Ser Asp
65 70 75 80Asp Glu Glu Lys Asn Pro Ser Lys Leu Leu Phe Leu Lys Thr Asn Lys
85 90 95Ser Asn Gly Asn Ile Thr Lys Asp Glu Pro Val Cys Ala Ile Lys Asn
100 105 110

s644PCT88.ST25

Glu Glu Glu Met Ala Pro Asp Gly Cys Glu Asp Ile Val Val Lys Ser
115 120 125

Phe Ser Glu Ser Gln Asn Lys Asp Glu Glu Phe Glu Lys Asp Lys Ile
130 135 140

Lys Met Lys Pro Lys Pro Arg Ile Leu Ser Ile Lys Ser Thr Ser Ser
145 150 155 160

Ala Glu Asn Asn Ser Leu Asp Thr Asp Asp His Phe Lys Pro Ser Pro
165 170 175

Trp Pro Arg Ser Met Leu Lys Lys Ser His Met Glu Glu Lys Asp
180 185 190

Gly Leu Glu Asp Lys Glu Thr Ala Leu Ser Glu Glu Leu Glu Leu His
195 200 205

Ser Ala Pro Ser Ser Leu Pro Thr Pro Asn Gly Ile Gln Leu Glu Ala
210 215 220

Glu Lys Lys Ala Phe Ser Glu Asn Leu Asp Pro Glu Asp Ser Cys Leu
225 230 235 240

Thr Ser Leu Ala Ser Ser Ser Leu Lys Gln Ile Leu Gly Asp Ser Phe
245 250 255

Ser Pro Gly Ser Glu Gly Asn Ala Ser Gly Lys Asp Pro Asn Glu Glu
260 265 270

Ile Thr Glu Asn His Asn Ser Leu Lys Ser Asp Glu Asn Lys Glu Asn
275 280 285

Ser Phe Ser Ala Asp His Val Thr Thr Ala Val Glu Lys Ser Lys Glu
290 295 300

Ser Gln Val Thr Ala Asp Asp Leu Glu Glu Lys Ala Lys Ala Glu
305 310 315 320

Leu Ile Met Asp Asp Asp Arg Thr Val Asp Pro Leu Leu Ser Lys Ser
325 330 335

Gln Ser Ile Leu Ile Ser Thr Ser Ala Thr Ala Ser Ser Lys Lys Thr
340 345 350

Ile Glu Asp Arg Asn Ile Lys Asn Lys Lys Ser Thr Asn Asn Arg Ala
355 360 365

Ser Ser Ala Ser Ala Arg Leu Met Thr Ser Glu Phe Leu Lys Lys Ser
370 375 380

S644PCT88.ST25

Ser Ser Lys Arg Arg Thr Pro Ser Thr Thr Thr Ser Ser His Tyr Leu
 385 390 395 400

Gly Thr Leu Lys Val Leu Asp Gln Lys Pro Ser Gln Lys Gln Ser Ile
 405 410 415

Glu Pro Asp Arg Ala Asp Asn Ile Arg Ala Ala Val Tyr Gln Glu Trp
 420 425 430

Leu Glu Lys Lys Asn Val Tyr Leu His Glu Met His Arg Ile Lys Arg
 435 440 445

Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn Glu Gln Lys Lys Ala Ala
 450 455 460

Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu Ala Trp Lys Ala Met Lys
 465 470 475 480

Glu Lys Glu Ala Lys Lys Ile Ala Ala Lys Lys Arg Leu Glu Glu Lys
 485 490 495

Asn Lys Lys Lys Thr Glu Glu Asn Ala Ala Arg Lys Gly Glu Ala
 500 505 510

Leu Gln Ala Phe Glu Lys Trp Lys Glu Lys Lys Met Glu Tyr Leu Lys
 515 520 525

Glu Lys Asn Arg Lys Glu Arg Glu Tyr Glu Arg Ala Lys Lys Gln Lys
 530 535 540

Glu Glu Glu Thr Val Ala Glu Lys Lys Lys Asp Asn Leu Thr Ala Val
 545 550 555 560

Glu Lys Trp Asn Glu Lys Lys Glu Ala Phe Phe Lys Gln Lys Lys
 565 570 575

Glu Lys Ile Asn Glu Lys Arg Lys Glu Glu Leu Lys Arg Ala Glu Lys
 580 585 590

Lys Asp Lys Asp Lys Gln Ala Ile Asn Glu Tyr Glu Lys Trp Leu Glu
 595 600 605

Asn Lys Glu Lys Gln Glu Arg Ile Glu Arg Lys Gln Lys Lys Arg His
 610 615 620

Ser Phe Leu Glu Ser Glu Ala Leu Pro Pro Trp Ser Pro Pro Ser Arg
 625 630 635 640

Thr Val Phe Ala Lys Val Phe
 645

s644PCT88.ST25

<210> 2

<211> 25

<212> PRT

<213> Homo sapiens

<400> 2

Met Ser Asp Glu Val Phe Ser Thr Thr Leu Ala Tyr Thr Lys Ser Pro
1 5 10 15

Lys Val Thr Lys Arg Thr Thr Phe Gln
20 25

<210> 3

<211> 28

<212> PRT

<213> Homo sapiens

<400> 3

Asp Glu Leu Ile Arg Ala Ile Thr Ala Arg Ser Ala Arg Gln Arg Ser
1 5 10 15

Ser Glu Tyr Ser Asp Asp Phe Asp Ser Asp Glu Ile
20 25

<210> 4

<211> 107

<212> PRT

<213> Homo sapiens

<400> 4

Val Ser Leu Gly Asp Phe Ser Asp Thr Ser Ala Asp Glu Asn Ser Val
1 5 10 15

Asn Lys Lys Met Asn Asp Phe His Ile Ser Asp Asp Glu Glu Lys Asn
20 25 30

Pro Ser Lys Leu Leu Phe Leu Lys Thr Asn Lys Ser Asn Gly Asn Ile
35 40 45

Thr Lys Asp Glu Pro Val Cys Ala Ile Lys Asn Glu Glu Glu Met Ala
50 55 60

s644PCT88.ST25

Pro Asp Gly Cys Glu Asp Ile Val Val Lys Ser Phe Ser Glu Ser Gln
65 70 75 80

Asn Lys Asp Glu Glu Phe Glu Lys Asp Lys Ile Lys Met Lys Pro Lys
85 90 95

Pro Arg Ile Leu Ser Ile Lys Ser Thr Ser Ser
100 105

<210> 5

<211> 76

<212> PRT

<213> Homo sapiens

<400> 5

Ala Glu Asn Asn Ser Leu Asp Thr Asp Asp His Phe Lys Pro Ser Pro
1 5 10 15

Trp Pro Arg Ser Met Leu Lys Lys Ser His Met Glu Glu Lys Asp
20 25 30

Gly Leu Glu Asp Lys Glu Thr Ala Leu Ser Glu Glu Leu Glu Leu His
35 40 45

Ser Ala Pro Ser Ser Leu Pro Thr Pro Asn Gly Ile Gln Leu Glu Ala
50 55 60

Glu Lys Lys Ala Phe Ser Glu Asn Leu Asp Pro Glu
65 70 75

<210> 6

<211> 31

<212> PRT

<213> Homo sapiens

<400> 6

Asp Ser Cys Leu Thr Ser Leu Ala Ser Ser Ser Leu Lys Gln Ile Leu
1 5 10 15

Gly Asp Ser Phe Ser Pro Gly Ser Glu Gly Asn Ala Ser Gly Lys
20 25 30

<210> 7

s644PCT88.ST25

<211> 83

<212> PRT

<213> Homo sapiens

<400> 7

Asp Pro Asn Glu Glu Ile Thr Glu Asn His Asn Ser Leu Lys Ser Asp
1 5 10 15

Glu Asn Lys Glu Asn Ser Phe Ser Ala Asp His Val Thr Thr Ala Val
20 25 30

Glu Lys Ser Lys Glu Ser Gln Val Thr Ala Asp Asp Leu Glu Glu
35 40 45

Lys Ala Lys Ala Glu Leu Ile Met Asp Asp Asp Arg Thr Val Asp Pro
50 55 60

Leu Leu Ser Lys Ser Gln Ser Ile Leu Ile Ser Thr Ser Ala Thr Ala
65 70 75 80

Ser Ser Lys

<210> 8

<211> 24

<212> PRT

<213> Homo sapiens

<400> 8

Lys Thr Ile Glu Asp Arg Asn Ile Lys Asn Lys Lys Ser Thr Asn Asn
1 5 10 15

Arg Ala Ser Ser Ala Ser Ala Arg
20

<210> 9

<211> 54

<212> PRT

<213> Homo sapiens

<400> 9

Leu Met Thr Ser Glu Phe Leu Lys Lys Ser Ser Ser Lys Arg Arg Thr
1 5 10 15

S644PCT88.ST25

Pro Ser Thr Thr Ser Ser His Tyr Leu Gly Thr Leu Lys Val Leu
20 25 30

Asp Gln Lys Pro Ser Gln Lys Gln Ser Ile Glu Pro Asp Arg Ala Asp
35 40 45

Asn Ile Arg Ala Ala Val
50

<210> 10

<211> 32

<212> PRT

<213> Homo sapiens

<400> 10

Tyr Gln Glu Trp Leu Glu Lys Lys Asn Val Tyr Leu His Glu Met His
1 5 10 15

Arg Ile Lys Arg Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn Glu Gln
20 25 30

<210> 11

<211> 54

<212> PRT

<213> Homo sapiens

<400> 11

Lys Lys Ala Ala Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu Ala Trp
1 5 10 15

Lys Ala Met Lys Glu Lys Glu Ala Lys Lys Ile Ala Ala Lys Lys Arg
20 25 30

Leu Glu Glu Lys Asn Lys Lys Thr Glu Glu Glu Asn Ala Ala Arg
35 40 45

Lys Gly Glu Ala Leu Gln
50

<210> 12

<211> 49

<212> PRT

s644PCT88.ST25

<213> Homo sapiens

<400> 12

Ala Phe Glu Lys Trp Lys Glu Lys Lys Met Glu Tyr Leu Lys Glu Lys
1 5 10 15

Asn Arg Lys Glu Arg Glu Tyr Glu Arg Ala Lys Lys Gln Lys Glu Glu
20 25 30

Glu Thr Val Ala Glu Lys Lys Asp Asn Leu Thr Ala Val Glu Lys
35 40 45

Trp

<210> 13

<211> 43

<212> PRT

<213> Homo sapiens

<400> 13

Asn Glu Lys Lys Glu Ala Phe Phe Lys Gln Lys Lys Lys Glu Lys Ile
1 5 10 15

Asn Glu Lys Arg Lys Glu Glu Leu Lys Arg Ala Glu Lys Lys Asp Lys
20 25 30

Asp Lys Gln Ala Ile Asn Glu Tyr Glu Lys Trp
35 40

<210> 14

<211> 41

<212> PRT

<213> Homo sapiens

<400> 14

Leu Glu Asn Lys Glu Lys Gln Glu Arg Ile Glu Arg Lys Gln Lys Lys
1 5 10 15

Arg His Ser Phe Leu Glu Ser Glu Ala Leu Pro Pro Trp Ser Pro Pro
20 25 30

Ser Arg Thr Val Phe Ala Lys Val Phe
35 40

S644PCT88.ST25

<210> 15
<211> 2575
<212> DNA
<213> Homo sapiens

<400> 15	
acttccttcg tctgggttgt tgccccagcg acacgttggg ccgaagagcg gtgttggta	60
cccgagagac ccggcggtgg ggaagtcact tcctccgaa gacgctgttt cctagcaacc	120
gccctccgcc tctgttatta gccccctcctc ctcgctcggt ccaggaccgg ctctgcggc	180
gccgccaggc ccagaccaag ctactatcag aagttgaatt ctaataatta gctattttat	240
aaaggtaacg agaaaaaaaata cactatgtct gatgaagttt ttagcaccac tttggcatat	300
acaagagtc caaaagttac caaaaagaact actttccagg atgagctaatt aagagcaatt	360
acagctcgct cagccagaca aaggagttct gaataactcag atgactttga cagtgtatgag	420
attgtttctt taggtgattt ttctgacact tcagcagatg aaaattcagt taataaaaaaa	480
atgaatgact ttcatatatc agatgatgaa gaaaagaatc cttcaaaact attgttttg	540
aaaaccaata aatcaaacgg taacataacc aaagatgagc cagtgtgtgc cataaaaat	600
gaagagggaaa tggcacctga tgggtgtgaa gacattgttg taaaatctt ctctgaatct	660
caaaataagg atgaggaatt tgaaaaagac aaaataaaaaa tgaaacctaa acccagaatt	720
ctttcaatta aaagcacatc ttccagcagaa aacaacagcc ttgacacaga tgatcacttt	780
aaaccatcac ttggccaag gagtatgtta aaaaagaaaa gtcacatgga ggagaaggat	840
ggactagaag ataaagaaac tgccctcagt gaagaattgg agttacattc tgcaccttct	900
tcccttccaa cgccgaatgg catacaatta gaagctgaga aaaaagcatt ctctgaaaac	960
cttgatcctg aggattcatg cttaacaagt cttagcatcat catcacttaa acaaattctt	1020
ggagattctt tttcaccagg atctgagggaa aacgcattcg gaaaagatcc aaatgaagaa	1080
atcactgaaa accataattc ttgaaatca gatgaaaata aagagaattc attttcagca	1140
gaccatgtga ctactgcagt tgagaaatcc aagggaaatgc aagtgactgc tgatgacctt	1200
gaagaagaaaa aggcaaaagc ggaactgatt atggatgatg acagaacagt tgatccacta	1260
ctatctaaat ctcagagtat cttaatatct accagtgc当地 cagcatcttca aagaaaaaca	1320
attgaagata gaaatataaa gaataaaaaag tcaacaaata atagagcatc cagtgc当地	1380
gccagattaa tgacctctga gtttttgaag aaatcttagtt ctaaaaggag aactccatcg	1440
acaactaccttctcaacta ttttagggact ttaaaagtct tggaccaaaa accttcacag	1500
aaacagagca tagaacctga tagagcagat aacataaggg cagctgttta tcaggagtgg	1560
tttagaaaaga aaaatgtata tttacatgaa atgcacagaa taaaaagaat tgaaagtgaa	1620
aacttaagga tccaaaatga acagaaaaaaa gctgctaaaaa gagaagaagc attagcatca	1680

s644PCT88.ST25

tttgaggcct ggaaggctat	gaaagaaaag	gaagcaaaga	aaatagctgc	caaaaagagg	1740
cttgaagaaa	aaaacaagaa	gaaaactgaa	gaagaaaatg	ctgcaagaaa	1800
ctacaagctt	ttgaaaaatg	gaaagagaaa	aagatggaat	atcttaaaga	1860
aaggagagag	aatatgaaag	agcaaagaaa	cagaaagagg	aggaaaactgt	1920
aagaaagata	attnaactgc	tgttgagaaa	tggaatgaaa	aaaaggaagc	1980
caaaagaaaa	aagaaaaat	aatgagaaa	agaaaggaag	aactgaaaag	2040
aaagataaag	ataaacaagc	tattaatgaa	tatgaaaaat	ggctggaaaa	2100
caagaaagaa	ttgaacgaaa	acagaagaaa	cgtcattcct	ttcttgaaag	2160
cctccgtgga	gccctccaag	cagaactgtg	ttcgcaaaag	tgtttgata	2220
ttacattatt	tggttattta	tcggttgcc	aatatttagcc	atagatttaa	2280
tatTTATGT	taggaaata	tatTTAATT	aaatGCCAGA	CACTCCTGCT	2340
gaaatacttt	ggaatgtaat	cagtgaaagc	atTTTTGA	ACTGTAGATA	2400
aacaaagacc	taataatcag	attgtttta	ccattaagat	acataagatt	2460
ctgataattc	ttatggtgga	gtgattcatg	atcttttca	ttaagctctg	2520
aagtatattt	aattccagta	ataaaaagga	aatcatctag	gtaccataaa	2575

<210> 16

<211> 29750

<212> DNA

<213> Homo sapiens

<400> 16	tctgggtgg	agttggcgg	gtcctgtctc	ctaggcaaca	gcacatgcac	acaagcgacc	60
aataatgagc	ccctctcaa	agacccagga	aggtgatgtc	acttccttcg	tctgggttgt		120
tgccccagcg	acacgttggg	ccgaagagcg	gtgttggta	ccc gagagac	ccggcggtgg		180
ggaagtca	tcctccgaa	gacgctgttt	cctagcaacc	gccctccgcc	tctgttatta		240
gcccctcc	ctcgctcggt	ccaggaccgg	ctctgcgggc	gccgccaggc	ccagaccaag		300
gtgagcagct	cctaccgat	gcttggctct	tgattctcag	ggtcgcggag	aactggccgc		360
ggcggtccgg	ggccgggaac	agaaagcggg	acctgggggc	catggggat	ccggacagag		420
accgcgttg	gacgtgcacg	ggcctggcgt	tcgctggtc	tcagcatacg	gcgcggtag		480
gagcggcgag	cacccggacg	tcacctggcc	tggtagggaa	cggaaccgg	ggcgcacaac		540
gctatggcg	gccctgccag	gcctctgctc	cgagtacggg	aaaccgcgt	tttatgcgg		600
ctcatcgca	aagcttcgtc	gttttgtctg	gctctttta	acactttgt	gagagaaaa		660
attggcttgc	aatacatctc	gctggctgtt	tgcgggttag	cattacgatc	tttttcttg		720
aatagcgctg	tatgcaaata	tatagataca	ttttttttt	ggtgggtgg	ctcataattt		780

S644PCT88.ST25

ttacggcgac gatcctttt atggcctttt aaataagacg tgacttattt tgaaggcaat	840
gttatacttt agaagagagg tgaaaaataa ggtgttctat tttaattggc agcattttgt	900
cgtattaact tgtaatcatt tatttcaga ctttttaagt agtgcaaaa ctattttagg	960
ataacttcca tttgaatttt tttaaacaag cttgttatga gaatttgcta tttctttaca	1020
agaacccttt taagtgaaga tgtagccaa tggtcatatc agatgtttt ctttgacctt	1080
tgtggggaga gtagaatcaa atgtaataaa ataaattctg aagcatgcga agtctgattt	1140
gttttgtata tttcagctac tatcagaagt tgaattctaa taatttagcta ttttataaag	1200
gtaacgagaa aaaatacact atgtctgatg aagtttttag caccactttg gcatatacaa	1260
agagtccaaa agttaccaa agaactactt tccaggtaaa gtatTTTtat ttggaatcat	1320
ttcacagtgt aaacactgta ttagatgggt tgaaattgggt gattctagaa cagtcctata	1380
taaagcaggg gtaaatctta tattactttt gaggtttgc acatgatcat gtttgggctc	1440
catccagtt tacaaactcc cctatatggt tttaagacta ccaaagttagc ctcaataacta	1500
gtttcctact aagttaaaag ttgaatcgca accttaaattt gccatttttataaaaaact	1560
ttttttctg ttgtacata atgttaagt tttttttct gttgagtcac tgcaattttt	1620
aactcagcct ctaagttgc aatattgatt gcatccattt ctgaaatatg ccgagacaaa	1680
agctcttaaa aataccaatt tctttcaaaa taccagttt taataaatta taatctaaat	1740
tgagccccctt cttatTTGTT accctccagc tctaattata acctgcaatt aatttggccc	1800
ataatgtgtg tctcctctag ttAAactgCG agctccatga ggaagggctc ttgtctgtga	1860
tgctctgcat tgagtatgag gcgtaaagtg ggtacatggc ataaagttag cttgcaggaa	1920
atatttggta gatgaatgaa acctaagttt gaaagcagtc gttaatcaag cattgtttgt	1980
ttaaagaatt acttgtgaat atgatacctc catgtttgga tggaaattga tttcagtatc	2040
tcatttcagg atgagctaat aagagcaatt acagctcgct cagccagaca aaggagttct	2100
gaatactcag atgactttga cagtgtatgag attggatgt gacagtatgg aaacgtgaac	2160
cactttctt cttttgttt ctttagttt gtatTTGCTT agccccccaa ccacccatcc	2220
cctcaatcac gtatgttaaa ataataccta agcatttact aatttttagat tttcaacttt	2280
ttaatttagta gaaagccact cttaattttc aggaagttgt atgatTTTCTT ttttttattt	2340
ttgtttgtt ttctgaatgt gtatacgaaa atataaatta attgtatggca ggtttgcagt	2400
aaaaggatgg ctgccagtgg taaaccacat tgaagaagac aggttcatct ttaagatcaa	2460
ccctaggagg tgctacagct agtttagtaac tagtcccaca gaactaaact tcggtgacca	2520
ttagaagtgc ttttataaag cttgtataa atcagatTTT tttggctgt gataaggggt	2580
aaattttaaa accacagact cttcgtttt catatatcag tactattata atttggtttc	2640
tcttagctat gtaaacatata taacattttt gtttcaggta taagcataca gaattctaaa	2700
cttgggtgttt ttgtttgttt gtttttgggtt ttgagatgga gtctcgctca gttgctcaag	2760
ctggagtgc gttggtgcaat ctcggctcac tgcaacctcc acctcccagg ttcaagtgtat	2820

s644PCT88.ST25

tctcctccctt cagcctcctg agtagctggg actacagggtg cccgccacca tgcccgctta	2880
atttttgtat ttttagtaga gatggggttt caccacatcg gccaggctgg tctcgaactc	2940
ctgaccttgt gatccgccccg cctcagcctc ccaaagtgtc gggattatag gtgtgagcca	3000
ccgcacccgg cctgggtgtt tattcttaa aatttggtaa ataattgtaa ttgatttctg	3060
taaaaccagt aataaccaca gttaaatcac tgctgtatag ttaacttagc atttctttag	3120
attcttagta aatctaatat tctgggtgtgg atggaattgt agttccaaaa ttttatgga	3180
aaaaatataa ttagtaatta ctaattaaat tcttccattt acaaatgttc ttgattttac	3240
atgaagaagt aatttgc当地 aaaaagttt acagtccata atctaattta aatgctacat	3300
gactgattgt tagggacctt tggatggctt tttccagagc aaacagtgtt tggttgggg	3360
gtaccctaca gacaacacaa taaatacatt ttgaataaat taatgaaatt ggaattttta	3420
tttcataaaat gttaatgaga cgtgcctgag ttagctgtgt ttttagagct gcaagtctat	3480
ttataaaata catttgtgcc tattcattgt tagaattttg ttttagctt ttaaggtaaa	3540
cttgattaa gttAACGtaa cttgacaat tttaaaaat actgttggaa acattttct	3600
tttccatttt tcagttctt taggtgattt ttctgacact tcagcagatg aaaattcagt	3660
taataaaaaa atgaatgact ttcataatc agatgatgaa gaaaagaatc cttcaaaact	3720
attgttttg aaaaccaata aatcaaacgg taacataacc aaagatgagc cagtgtgtgc	3780
catcaaaaat gaagaggaaa tggcacctga tgggtgtgaa gacattgtt taaaatctt	3840
ctctgaatct caaaataagg atgaggaatt tgaaaaagac aaaataaaaa tgaaacctaa	3900
acccagaatt cttaaattttaa aaagcacatc ttcaggtaat ttgttaggat tactgttaatt	3960
gcatttcttga gaagtttatt ttaagataat cagtc当地aaa attttatat ggtagctagt	4020
atataatttaa gaaaaaaaaa cagacttaac ttccattttt cagacctgtt gtatTTgtc	4080
taacttcaat ttacagacc tttttgtattt tgtcttaactt caattttaca gacctgtgt	4140
attttgc当地t gcatctaggc ttttgcctga tagaaagcca aagcacaaa ccaaagcacc	4200
tttagtcatc catagcatcc atagctgtgg atctccagac accttagaccc gtgagcttca	4260
ttttgtttt taggtgtgga actggatgg aatgctgtct aatccctctc acactccaaa	4320
gattagagtt acagcaatat tgagactaat ctttcaaca gtctttgcca taccacatt	4380
gtgccagaaa atttcttga catttgc当地 tttgaaggat gagttatgtt attgctgctg	4440
ttgtttgtt aagcatccag gcactccttta agagaatctc catttgc当地 ctgtattgcc	4500
tatgaaaatc tactaagatt cagtttcca aaggaaagtt cctgggtgtga tctgggatta	4560
cagttagttc tgcccacaat ttactgaat tttaagcata aaggaacaaa gatagaatga	4620
aacggagacc aagtccgtc acataccctg ggccaccatt catgaacttgc tatatgcaag	4680
gttaaggatt tttttttt catttttgtt atttataaa ggaatttatta gttgatgtt	4740
accttcataa aaatccctt gcataatcatc agtaaataca gtgctggtaa atatttcata	4800
ctttgc当地at tagataccag tggtaacgtc agacaaaact ttatTCAGG catgtattgg	4860

S644PCT88.ST25

ggaaactgctc	ctttcttcct	gaccccacaa	tctcatttaac	tttgaatga	gcaaaggatg	4920
taagcagagc	aaagaacact	agaataatat	ccaggacact	ggggaaagg	cctctgtata	4980
ttatatatga	cttcagcaaa	taagtttaagc	ttcagtatcc	tcatgatgag	gaagctaaaa	5040
ataaccctct	ttcttattcct	gcaaaattgt	gagagtttat	tgaagtgcatt	ctcataaaact	5100
ataaaaaaact	acaaaaatgc	aaacagatgc	ataatgaaac	aattaacttg	ttaaaatgt	5160
ccttctaagt	atagttagtgc	aatcaatgc	tggagagaag	aggaacataa	ttgaacttcg	5220
ttattaagaa	aatgcgagca	tatatacgaa	ctaaaaattt	gtctgagaca	ggtgatgt	5280
tataattttaga	agtttatgtt	agataatcag	gaaagcaata	atccacctat	ttcatacctt	5340
aaaaaaaaaa	aaaacctgtg	gtgggttaca	atgaataaga	aaatactgt	ttttaaccac	5400
aaggtggcat	caggatccta	aatgctctac	ttatataatgc	aatgttatat	tcagtacgt	5460
taatataaaa	ataattacct	aaataggtaa	ttgtatacat	tgattaccaa	aaaaagcgct	5520
tttcttaaag	tataggcatt	ttttttctt	tttggaaact	tgacagtact	tctgaaagt	5580
gaatttttgt	agaaaatata	ttaaagttgt	cattctcagg	ttcttcaggt	tgaaaagtaa	5640
aaattgaggc	tagtggcct	aagataatat	ctggcatata	taataagtat	ttaaatgt	5700
aaattaatat	atgaatgatt	tatcttgc	agagggata	tggcatgt	gtttatcctc	5760
taaattcttt	gactttttt	tttctgtac	aggttggaa	ctcaatgttt	ttaatgtgt	5820
gagatattgc	tgagtagcaa	gtaatgcttt	atgaaactat	tagagcttgc	aggtttctc	5880
tgtccctgct	tgtctttgt	aaaaagtata	ataaccagac	tttatagtca	ctactgaagt	5940
gacagttgct	ctataaagt	aaagtatttt	tcacaggata	tgttttatt	ttaatactaa	6000
catgactgaa	atcatgaact	ttggagtcag	gatgcttc	ctttaatctg	agatctgcag	6060
cctgctagag	tttgtgactt	tggcatgag	acctcttgc	tctcattttt	ttcatcttta	6120
aaaacgggat	aatagttgcc	tgcctctagg	agtttgaggc	aattaaatgt	gttcacat	6180
ttgaagtgc	tagaatagta	ctggcataaa	tttagcactc	tataaatgtt	ctgattattc	6240
attttattat	ttagcgttt	tttataaaca	tgctcagcag	gtataaaagta	tcagtcgt	6300
gggatgcgt	agttctagag	atctgctgt	cattgtgcct	atagtttaca	gtactgtctt	6360
ttgcactgaa	tgtttaaga	aggttagatct	catgtttgtt	cttaccacaa	taataaaaaaa	6420
aattgactca	acaccttctt	tcaggcatta	tataatattc	tgctttaact	gaggctcaaa	6480
agacatgca	gcatttgc	ggaggagaag	caggaagtgg	atattctagg	cagggggatc	6540
agcttaggt	aaggtatgtt	agcaggaggg	attggaggg	ttgtggatgt	tgtgcgt	6600
aactgttagc	ccagcatttc	agaaacacag	atgacaaaat	ggctgttagat	aaggcagt	6660
aggacaaaac	cataaaatcc	gttttatgtt	gtttaaaggc	agtttagctt	ttattctgt	6720
ggattggatc	atggggagcc	attgaataat	ttttagaaaa	ggagtgtatgt	gatctgat	6780
ggattttgt	aatatcatgg	aagcgtgt	ctaggaaaga	gtggataagg	acccgacagc	6840
aggatgttag	aaagtggaaat	aatgagata	tttggcaatt	agaattgtata	ggatatttgc	6900

s644PCT88.ST25

atactctgga	tttagggat	aatagaggga	ggaatctaga	gcccttggat	ttggggttga	6960
acatttggct	ggagtttagg	atgttagctaa	aattgtcagc	tacttataat	aataccaatt	7020
tggtatggtt	gtggaatctt	ctggcagaat	ccataagccc	attttttagt	aatgggagg	7080
aagatgttaa	ttagaccaat	tttgaagttg	agaaaaatgc	atttgtagaa	caatagaaac	7140
ataaatatgt	atagcaggt	aatgcaggc	aaaaaatata	tacatggaaa	gtcttcccat	7200
tgttcgaat	actggatgc	aatcagcatt	tgattcttga	tttaaactta	gaagtaatgg	7260
aaagagtcaa	attttaataa	atgctaaaga	agtttatgg	actcagaaca	atthaactcat	7320
aaaagattcc	ttcctcta	gagagttgc	actcctatcc	cttggatgcc	aacatcatca	7380
tcttgcct	tataatagca	cttataatct	tagtaatcta	gtcttgtaat	tttgtttaga	7440
aaaatcaacc	tgtaaagtac	ctggacaggt	ccattgccgc	tttggatt	atgaggttt	7500
gtaacgtgt	cagggcttgg	tactcaaagg	cttgatggat	gagcctcctc	attttatagt	7560
ggtagaaact	ggggcaagat	tttggttgt	ttttttat	ttaacattt	tttttaata	7620
ttataagagt	tcacaatgtt	gaagagttaa	cttctgtga	ctggttactt	tcaggatgac	7680
aactgtttct	ttactttgtt	tttttttgt	tgttggttt	ttttggttt	ttttttttt	7740
tttagatggat	tttgctctt	attacccagg	ctggagtgc	gtgggtgtat	ctcgatctcg	7800
gctcactgca	acctcagact	cctgggttca	agcaatcctc	ctgcctcagt	ctcctgagta	7860
gctgggatta	caggcacg	ctactaagcc	cggctaattt	ttttgtat	tttagtagaga	7920
cagggttca	ccgtgttagc	caggctggc	tcgaactcct	gacctcatga	tctgcccacc	7980
tcggcctccc	aacgtgctgg	gattacaggc	gtgagtcacc	gctcccaaca	tgtcgggatc	8040
acaggcgt	gccaccgcgt	ccggcctgat	tattaaccat	catttattt	tgccttacta	8100
gagctctgt	tagagaagag	ttgtggcctt	catctggact	cttcaggaca	gagaacaaag	8160
gggcataggc	acaggaggga	agtatggtag	cacccagaga	gataaaaaaa	gccatggtca	8220
ttttttata	cacacactt	aagcatttta	ttttcagca	aaaaacaaca	gccttgacac	8280
agatgatcac	tttaaaccat	cacctcgccc	aaggagtatg	ttgaaaaaga	aaagtcacat	8340
ggaggagaag	gatggactag	aagataaaga	aactgccctc	agtgaagaat	tggagttaca	8400
ttctgcacct	tctcccttc	caacgccc	tggcatacaa	ttagaagctg	agaaaaaaagc	8460
attctctgaa	aaccttgatc	ctgaggtag	cactaccact	aaactgttga	attgtgttct	8520
tgaatttat	ctttttatc	tgattatgaa	aaagagaagg	agagaatgaa	tttgtgtgcg	8580
tgtgtgtgt	ttttacatac	tttcttctgc	aactgataag	gaaataattt	ttaaaaatac	8640
actgtattcc	accgagtcta	aaactgcac	aattgtaa	cgtgcattat	ttttacatac	8700
cactaaggaa	gaagggaaatg	catccaatta	aactataaca	caccagtat	tgttagagtt	8760
atccagttt	agagaaagta	aatgtcaa	aagtgttgct	tttctgaatc	tatataatag	8820
tgtttatctt	taataat	ttaaattt	gtatcttga	attatgtat	ttatggctaa	8880
gaacaatata	gtcagtgtca	ttttat	ttgat	ttactcaaca	aatgtgtgtt	8940

s644PCT88.ST25

gaatgttcat ggcactcttc tgtgttcttt	gggttatgtt ccaatagcat taaatgtggc	9000
cttcagggtt tccatcaggg aatttactat	gcattgttat taagggagaa cacttcgttt	9060
ttctctttgt atttcactat gagaagcaaa	ctgtcccttc tgaacatttc agaaggaaaa	9120
agtacaggaa gaacatttct tccccataat	ctgcttgggc agattaggaa actgcattgcc	9180
acctggccaa gcttcttct ttttctcatc	gcttgtctgc agtgttggtg cttaaggatc	9240
tgctctctgg gaggtgagggc agaagggtct	gagaggagct cttttgtgca atgactaaat	9300
gggggaatcc ccctaattca gacttggaaat	atttaggaagc acaataggct accaattcaa	9360
atcttgttct gcagttgagc tttaccagta	aagctgacaa tttgatatac gcctaactga	9420
caccaccatg ctgtttctta atttgttctg	aaaaccagaa gaagaaacccc aagcaaatac	9480
tttatattta agaaaattat ctgatccatt	gaatattgtg cttagttctt gtagctgctg	9540
taacaaatttgc acacaaactg gttaacttaa	aacaacagaa atgtattctc ttagttctgg	9600
aggtcagaag tccaagatca aggtgtttgc	agggccattt tcctctgaag gcatcacgga	9660
agaatcccttc cttgcctctt ccagcttctt	tcttagtggtt gccagcagtc catggcattc	9720
cttggcttgt agctggcttg tagctgcattc	attcccttct ctgccttcat cccatgtggc	9780
cttctccct gtgtttctc tgcatgtctg	tgtctcttct ttctcttaaa aaaagacacc	9840
aggcatttggaa ttttagggccc accctaattt	agtgtgtcct catcttatct atttaaagct	9900
gtaaacacct tatttcctaa gaaagtcgta	ttttgagggtt ctggatgaac atgaattttg	9960
gggcatttaat gttcgatgt taaaccttagc	attcccccggaa taaactctgg ttagtcatgg	10020
tgtgatattt tattgtggaa tttgtatgt	taaaattgtg ttaagggtttt catcttatatt	10080
tatgaagtct attggctctgt aatttttttc	ttataatgtt accatcaggc ttgggtatca	10140
aatgagttgg ggagtgtctt ttcttcattt	tataaaagtt tggatcatt attttcttaa	10200
atgagaggat tcaccagtagc aattatctgg	gcctggaaatt ttctgtgtgg agacatctt	10260
ggcattacat ttgattttt aaataggat	ttcagactc acattttctg ttttgcagt	10320
ttggtaattt tttttttttt	tttgcattttt gatatgttga gtttataaac	10380
agagttgttc acgatagtcc ctcatcttt	tgatgactag gattatcatg acatttcatt	10440
tttatttcta acatatataa tttgtgtttt	gtgtcttctg tgctaaatct tgataggcat	10500
tgcttagttt tattaaacgt ttttaagaac	cacttcggct ttgtcatatg ttggtgcaaa	10560
agtaatttgc gttttggcca ttactttcaa	tgacaaaaac cgcaatcatt ttgcaccaac	10620
ctaataattt tctctattgt ttgtttaatt	gattttcagt attatttcag tattattcag	10680
tattatttct ttactttct tttttttttt	ttgagacaga gtctcgatct atcgcccagg	10740
ctggagtgca gtggtgcaat cccagctcac	tgcaagctct gcctcccagg ttcactccat	10800
tctccctgctt cagcctcccg agtagctggg	actacaggca cccaccacca tgcctggcta	10860
atttttgtat ttttagtaga gacggggttt	caccgcgtta gccaggatgg tctcgatctc	10920
ctgacatcgt gatccaccca cctcggcctc	ccaagggtgtt gggattacag gcgtgagcca	10980

S644PCT88.ST25

cggcgccctgg cctctttac ttcttttgg ttaattgc ttatcttag atttgaat 11040
 ttctcattc attttaaga tttcgtat ttctgctaaa cctgttggaa ggtgtaaact 11100
 ttcttcttg tactgctta gtggccccga tttttgatg ccttttattt ttattatcat 11160
 ttcttaaat atatattta acttcccttg tgatctctg tttaaaaat ttatttttt 11220
 agttaaaaaa taataattgt acatgggta catagtattt tttcgataca tataatata 11280
 agtgatcatt gtgatctctt tttgaccag ttggttattt tatggtgatt tattttattt 11340
 tcaaatactt gtttttctc tagatatact tttgatgtt attataagtt aattttgtt 11400
 tagtctagag aatgtatctt acatgattt aaattttaa aaatttattt tattttct 11460
 aaatggcca gcttttagtgt atcttgcattt gcatctgcaa agtagatgtg 11520
 ttctccagggt gttgaatata atgttgcattt atttaagttt ggtcaacatg gttggtaata 11580
 tcattcagat cttcttatac cttactgatt tttcatccaa tttgttacc cgttaccaac 11640
 tttagggtat taaaatatcc agttatgttt gtgggttgc ttatacttct cttagttct 11700
 gtcagtattt tataactttg ttatcaggca catacacatt tattttattt atgttttgag 11760
 cattatgaaa cgtctctacc tctggtaata ttccttcct tatctttagt attgttttgt 11820
 gtaataacttc agcttctta tgacaagtgt ttccatggta tatgctttct atctttttc 11880
 tttcaaacta attctgtctt ttcatgttattt tgaatctctt acaataagag tttgggtca 11940
 ctttttattt aagtctgaca atctatgcct tttatgttag tgtttagtcc atttatgaat 12000
 gttttgtcca tttatgttattt atactgctat gattggattt aggagcaatt tggtctt 12060
 tattttctat ttatctgttt tttaaaatattt ttgtttttat tggttttct ctgttactcc 12120
 tttcttgccc ttttttggg agataatcat gaatcttta gtttttattt attattgacc 12180
 ttttatctat atttgggttgc attgttatttcc tcagagttga tcagtggtt accaaatata 12240
 tctgaaaatt atcacaatct atttagaatt gatattgtat tgttcacat ttgatctaga 12300
 aaccttgaa taatatagtt ccatatactc cctcatccat tgtgttattt tcatatatta 12360
 tatctacata tcctataatc cccacaatag agttataact tttctttaa gagcccttc 12420
 agtttttgtt attagacttt taaaatattt aagaaggcta gaataaatat atattatata 12480
 tctactgtat tatataattgtt atatattata gataacattc tattgttattt tataatata 12540
 atatattgtt agacaatatc tatataatagg taatataat tctattctt tatattatata 12600
 agatataaa catctatata atctatttattt agatattaca tatctataaa tacatataca 12660
 atttcttaggg atcttcattt cttctgttag attcagatta ccattttgtg tcctgtcagt 12720
 cttacaaact tattttacat ttctgttattt acagggttac tagtgttgc ttttctcag 12780
 tctttgttt tctaaaatgtt tttgtctcat ctttgggttcc aaatgggtt ggtgttgcattt 12840
 gtattcttct tgcattacat ttgccttctt ctacccatgtt ctctttatag gtttccattt 12900
 ttattggcct ctctgttattt catttcatttc attgttgcattt ctatataatg tttgttgcattt 12960
 gtctgaatgc tgcattacat ttactcaag attgtgggtt ttatctttt attacagcaa 13020

S644PCT88.ST25

tttgactgca tggtgcctgg gtctagctt ctttatgttt attctgcttg acgtttgtt 13080
 agctttccaa acctataagc tgatactgtc tgtgaaatgg gaagattgtt atttcccacc 13140
 ctattttca tcctctccctt ttggtactgt agttacacat gcattgaaat ttgtgctata 13200
 tctcactgat ctctgagatt ctgtttatat ttcttaaatac tttttccctc tttgtttta 13260
 agattgaata acttgttatta cttagtcttc acgtttacag attgtggtcc ggagaatgta 13320
 tctttatga tttcaaattt tattaaatta ttttggggg ttttaatggc ccagcaaaag 13380
 ggtatgtcgt gagagttcca tttgcagttg caaagtatgt gtgtttcca ggtgaatttt 13440
 ttatttcact tattgtgggt ttcaacttca gattttctat ttgttatttt tctgttttt 13500
 aatataaaaat cccccatctt ttcagccatc atgcataat tttccccaaa gtgctgaac 13560
 atatttataat tagcttattt aaagtccttg tctgctaact ctaaaacgtg agtcatctc 13620
 gggttgggttc ctattgacca ttctctgtt ttttatttt ttttttaaat aagtgtcacc 13680
 attttctgtt tcttttagtga ctttgattt aataccgggt gttctgaatg atatttgtt 13740
 gagattctgt attctttat gtcccttcaa acatatttc tagcaagtgg atatcatggc 13800
 tggacacaaa ttcccaatcc ttttctcct gcagtggttca tcagctgaaa tttctgctt 13860
 attctttca gtttctagct tctatgttt tacaggatcc tctgaggctt cccttatgcc 13920
 acaaataagag gtggtaaagg ttttgggttga atttcatatg cagattttgt ggtcaactgtc 13980
 ctctgctatt ttccacatc ttattggctg atctgatggt cctagactca gtccccctgtt 14040
 ccctcaagtc attccaccaa ggctgttagcc ttcttattact tgagctgcat agactggaga 14100
 atgccttctg gcaaaaagct actaatttgc agatctcctc aggtgaagct ttatcttca 14160
 gggtagactc cagtgtctca gcacttcttc cattttctca aatgtttctt ctccattgtt 14220
 tttgacatat aatttcctt gcacccataa aatactgcgg agaaagaaaa tttaagtatt 14280
 tgtacaacaa agttgaactt cctacattgt aatatcatta cctttaggct agatgattct 14340
 atgaagaaaat gtttacctt gatagacaaa tataatttt tcatatcaga tagaattttc 14400
 agaattttga gggaaaactca agtgcatttca atctatgtgc ttttcctatc taaaatattt 14460
 ggaagtagcg gcttacttga ttttattaaa tgctttcatt tggtataacta gtaatattt 14520
 cttggacta aagtattttt cctgtcttct ttatgctttc cttcaaagga taattttagg 14580
 aagagctatc aaaatcaaattt cttggccttca aatatttata agaaatgttta ttattaaat 14640
 ataggagttt tgaaaattgg taaaaataa atagagaggt ggtggtagtt aaagaacttg 14700
 aataactctt tcagtgcaccc ctttaatgtt ccaagacatc aaggcttcaa agttaaagcat 14760
 gcttacctcc attggcttgc cacactttgc gtttcagcaa caaatgccta aataatgcag 14820
 atttcagagt tatgcactat ttcaattttgt agtttaata atgctattgt tccctataat 14880
 gttaattttt aaacttatgtt ggcaaattgtt ttttttttgc gggaaaacagg attcatgctt 14940
 aacaagtcttta gcatcatcat cacttaaaca aattcttggaa gattttttt caccaggatc 15000
 tgaggaaac gcatctggaa aagggtggta tatctaataa ttatatctt tatgtgaact 15060

s644PCT88, ST25

ctgtactact tagactcctg tttgttaagag aaataatact ttgtatagtt ataagagaaa	15120
tatatgttt tatgtgtttg agtttaatc ctgactatgt agttaactaa ctgtatttt	15180
ggatgcagaa cttaatctct cagtgcctca atttccctaa gtttatattat ttgtctcata	15240
aggttattgt gaaaattaag tgatatagtg catttttagcc attagcctag ttaatagccc	15300
aagtggagtg agcacttaag gtaaactact gttatgtatg tggctgtg atattctgca	15360
ggacaacata atagcttagt ggaattttaa agtgagacta agctagattc caatacaggc	15420
acaattacat aagcaaagta actaaccttt ctgaccctgt atgttgcattt taaaaatggg	15480
taaaataaga gtaatttgcc ttatagggtg ttgtaaat taaacatgtt aagcatttac	15540
agcaataccca tagtaagcac ttgggtgtat atgtgaattt ttaacataat ttctttctt	15600
agtgatacgt agcttaatga aacctaaaag acatagctat ttcttaggtct gagatgtgt	15660
atgaacattt tagtgcttac tatgtatgtat cattttgtc attttacaga tgagaaaagc	15720
tgaagtgcag tgacttaggg aaacataccca aaggtcagtg atgaaaccat agttaaatct	15780
ttagttccaa agttcttgg ctttcactg aacagattaa cagctccaaa gaatccaata	15840
gtgaatttgag tgattttaaag cccatgttac ctcaaaacaa attccaaaaa aatggtcata	15900
atgaaaccaa cagaattaag acttttcaca gtaaagattc aggtttagct gcaaggtgga	15960
cgttggtaga actgaaagtt ggtgatccc ttccaaaatg tggtaaaatc agaataatgt	16020
aagcaattct ataaatgcaa aactgaatct tcttatgcc aagcttgagc ctgtttctt	16080
gagcactgag aggataagca ataggcttgt ctttattgcc ccttatggta tcagaggaag	16140
tactacatct tggtgagatg aaactcacta gagactgtgt aaaattgcat taattcttgg	16200
ttcttctgc agctatacaa ttcaacaatt gtactactag taactgttagt agcctagaga	16260
ggtgtgacac cttcttatgc agcgtgttgt tccagctaag aaactcaggc tttagagttt	16320
aacaaatatt gtcatctcac ttacttgggt tttatataatca caagctctt tgacatgtcg	16380
ttgttttagg gtatgtttc cattctgttt attaatatgc tttttctt agtactatgt	16440
ttgttaagtg cttcattatgt taagcctaga ctatttttt ttgtaaatca ctttcgaaaa	16500
gagtttatgc aagtttaata tgataacttt tttcatatt ttgcaagaaa aaagagttt	16560
tagatagtcc tcattttaaa gaaagcaaataat gaaatcaatgtt tttacctttaatc	16620
gggggtttta atgctattac tctgtctcaa aatagatcca aatgaagaaa tcactgaaaa	16680
ccataattcc ttgaaatcag atgaaaataa agagaattca ttttcagcag accatgtgac	16740
tactgcagtt gagaaatcca aggaaagtca agtgactgct gatgacccctt aagaagaaaa	16800
ggcaaaaagcg gaaactgatta tggatgtga cagaacagtt gatccactac tatctaaatc	16860
tcagagtatc ttaatatctt ccagtgcac agcatcttca aaggtattt taaaattca	16920
tactttcat actacagtt aaaacttgaa atagaacttt aagaaattttt atcttctgtg	16980
ttatataactt ctgaattacc agtggaaaat ttatctttt atagtgtat tttttgtca	17040
catggttctt acttaatcca ataaaatttta acttttaagga aagtttgtag tgaatataat	17100

s644PCT88.ST25

gaaaccaggat gttaaaaat tatcagaggt gtgtgatcat aataacttt taaatgtctc 17160
 agaaatgcattactcatagtg tatatatatttc cataggtctt catattttaa aaataact 17220
 gtctggata atttctgaga ttttaaatta gagttatgtt ttggatatt gtttaaaac 17280
 gtgttaacaa ttttaacaaa aatcttaaag aatgtttat caacagtttcaacatctg 17340
 tgcttctta aatagatgg ttatcatcag gaacattagt attattattc gtatttgatc 17400
 cttgccttt atttcctaattttcaaaata atgaactggt gccctggcaa cctccagagg 17460
 ttagtggatc gctttgttt ttctttttc aattcatgta aatttaatgg ttacaagtgc 17520
 tttttgtta catggatata ttgtgttagt gtaaagtca accttttagta taaactaaaa 17580
 tgtacattgt acccattaag taatttctca tcccgcacct ccctctcacc tttcttagtc 17640
 tccattatct attattccat accctatata catgtgtaca cattattnnctgacttg 17700
 taagtggagaa catgtaccat ttgactttct gtttctgatt tatttactt aaggtaatag 17760
 cctccagttc catccatgtt gtaaaagata ttatttctt tctgtgtggc tgaatagtat 17820
 tcctgtgtgt gtgtgtgtgt gtgtgtgtgt atacacattt tctttataca 17880
 atcatatgtt gatgtacact taggttgatt ccatatctt gctattgtga ctgtggtgt 17940
 gataaacatg agtgcaggta tctttttat ataatgattt atttcctt tggcagatac 18000
 tcacagtggg gttgctggat tgagtggtag ttctatattt agttccttaa gaaatcccc 18060
 aactattttc cataaagatt gtactaattt acattcttac caagagtata caagcattcc 18120
 cttttctctg ttttctcacc aacatctgtt actttttaa ctttttaata atagctaaat 18180
 attctgacta gtataatata tctcactgtg gtttaattt gtgtttctct gatgatttagt 18240
 gatggtaac attttttttc atgtttcttg gccacttgta tgtcttctt tcaaaaagtc 18300
 tattcatgtt ttttgccttc ttttagtgg ggttatttgt ttttggatgt tggttttag 18360
 ggaaacatta ttattataac cttaagaaac agatatgtaa tatgttaggat tacttgccc 18420
 tacattaaat tgtgcctgag tgctatactt taaaaatttta tggtgtagca tttcagtct 18480
 ttgtttctcc tgaatttgcattatctctt gtagctgaa ttagctgca gctctgtgt 18540
 tttattatca gcgaaagaaa acagggctag ctgaaaattt gtgtttgagc aatactttt 18600
 taacataaaa tacaagcttt tcttaaaattt gatgaaggag gttcattaag ccatgttcca 18660
 ggttatcat ctttagctaa ttcttttagg aaaaaaacac tactgctaag ttagggatgt 18720
 gtttattatg tctgtgtctt cactttacca ctgcacccca tcagtcgtg taaagtagaa 18780
 aagttgttcc taaaagaag aaaggatatt ccggagtttta tagacaggat tggatgt 18840
 ctaatagagg caattctaaa ttagaaçagg catttcataat gtaacaagta aggttgcac 18900
 ttgtttcttt tgactggacc cttggcctca ttcttactct ctactgaatg acctttctt 18960
 aacagaaata taatcattct ccattaaagt cttttgttg gtttctcatc acaagaattc 19020
 catccagact cctcatcgct gccttagtgc ttcacctggt tcttccctga ccacgtcttc 19080
 ctccgccttc cctgccattc actatgttcc agctccattc accttttctt gttttcaga 19140

S644PCT88.ST25

gataacaggt tccgtccctt ctcaggcttt tacccacttg ctgtttcttt ctttcataga 19200
 cctttcggtg ggccctttgc actcttagct ctgatgtcag cccctcagga cagccttccc 19260
 tgaccaacct cttaaagca gctcctcagc cccactctag tcattctctg tcactgcaca 19320
 ctatttatg tccttcatga gccatgtttg cttatatatt tatttttgt catccgtctc 19380
 tagaatttaa tattcttaag ggcattttat tcactgattt gctcccaatt tctactgtgt 19440
 ttgacacata gtagatgct aaagaatagt gatttactgg cagtttgct tctaagccta 19500
 aaaaggatag ttgtcatgaa taaatcatct ttggcatttt ctgtttaata gaaaacaatt 19560
 gaagatagaa atataaagaa taaaaagtca acaaataata gagcatccag tgcatctgcc 19620
 aggttaataaa gttaccaata tttgtcattt atgggcttgc attctagcaa agctagttt 19680
 aatttaactt tcataaaagta aatttcattt ggtgttactg tattttcttt ttatttccat 19740
 ttcataaaat gaaagtagtt aacttcatga taaaacccct tgggtgatga tattatttga 19800
 aataaagtaa ttataaaaaa gtaagtctat tactgattgt ttttagtgcct ggaatgttta 19860
 tgcaataacct ttgctctcca ggatcgtcct aggaatattt ttcttcttc ttaatgtcag 19920
 tgattaggga ttctttgtgc tccagactgc ttctggaata gagttcttt ctcctacttt 19980
 tcctgagaca agcaatataa aatggtaata aagctgaagt ctagcaatga tacttattca 20040
 ttatcaagta tcattgtcta acatgagaaa ttgtactgaa agcttcaga atctatgaac 20100
 taagtaggtt tattaaaatg attatctgta tagtttcatt cacaccaatg ataatgaatg 20160
 cctaactcat aagtgctaata caaaaacctt ctgaatcttt aaaattatcg ttagtcaaata 20220
 tattcattaat caaataaaac agagctagca agcttttct gtaaaatggcc agtttagtgca 20280
 tattttaggc tttgttaggcg atacagtctg tatttggact actcatttct gctattttaa 20340
 cagggaaagca gccacaggca aaacttaaca tgaatgatta cagctatggc gcaataaaact 20400
 ttgtatataa aaaccaatgg ctggccaaat tttcccacca atccctgata tagatagtac 20460
 tattcttct aattttatat ttggaatgct tcatgtaaaca aatgtatgaa agaaaatatt 20520
 aaaagagtga ttataaccctt ctgtattgtt tttccatgt aacttgagaa gtggccata 20580
 tttcttaagt ttcttaattac aaatattaa aaagagcaat cattttaaag ctatataact 20640
 taaagttata aaatttaaat tatgttgaag gggacatatt taagttatgt ccccttctac 20700
 ataatttaat attctttgtt tactaagact gtacattttt cctacatcat tttcaaagta 20760
 attataattt gttaaatttat aatgttagttt ccaattttt ttttggatgt gagtctcact 20820
 ctgttgctca ggctggagtt cagttggcatg atctctgctc actgcaacct ctgcctcctg 20880
 ggctcaagct atccctccac ctcagcctcc aggtagctt tgactacagg catgtgccac 20940
 cacgccagct aatttttgtt attttggta gagacagggt ttcaccatgt tgcccaggct 21000
 ggtcaacagc ccaacaggat gagctcaagt catccaccca ctttggccctt ccaaagtgc 21060
 gggattacag gtgtgagcca tcatgcctgg ccagtttca aatattatac gtgcattttc 21120
 taacagatct ctcttctacc aaatgcaattt gtaatatttt gtcttgattt atttggatct 21180

s644PCT88.ST25

tttcagatta atgacacctg agttttgaa gaaatctgt tctaaaaggga gaactccatc 21240
 gacaactacc tcttctcaact atttagggac tttaaaaagtc ttggaccaaa aacccatcaca 21300
 gaaacagagc atagaacctg atagagcaga taacataagg gcagctgttt atcaggtaaa 21360
 aaaggaaaat atttttaaga gaagaagaat gatcactttc ataaggctac actgtttata 21420
 aagaataaaag taatcctgtat agaaaaatgtat ggtttaatac ttaaaatttat tgagaaagag 21480
 tttcctttta atacatgagt aatcatatTT tactaaatta tttgcttcca cactttgcatt 21540
 aactgaccat agttgtttt aaagaaagaa tatgccattg caatttatacg aaatacagca 21600
 caagccaaaaa cattgtaaag tctatatacg ttttcatttt tttcttcttg aagtttata 21660
 gaacaaaaagg agttattatg aacaaaaagt tattaaattt tttcttccct gagatgttgt 21720
 taggcgtaca taggaaaaag attgtattaa tttattcaca attctaaaag tcttttttg 21780
 tcttttttag agtagaatag tatacttttag aaaattgtac atgtgaattt cagagaaaaat 21840
 gttaatataa agaattctaa ttcacttaag aaattttaaa tattatatacg cctttttctt 21900
 gttcttatag gagtggttag aaaagaaaaa tgtgtattta catgaaatgc acagaataaa 21960
 aagaattgaa agtggaaact taaggatcca aatgaacag gtattctgac atatagaagt 22020
 aaaaatgttt tggattttta tttcgtaaa atatcccgtat atatataact tttctaaatc 22080
 agcttttaa atggcaaaat aacttgcata ttaaagaaat gatttccgggt tttacttctg 22140
 ttttacttta tacatttttag tttgatataa ctgtttaca tgaaaacaga ttttaatttt 22200
 gtatatgtat aggatagctt tgttcctgct gattatgaag ttattattgt ttatgagcac 22260
 ctaattcact tttaaaagtt gatttcattt agaacttaac caagaaggcc aggtactgtg 22320
 gctcatgcct gtaatcccag cactttggga ggccaaggca gatgggattc cttgaggct 22380
 ggagttcgac accagcctgg gcaatgtggt gaaacccat ctctactaaa aataaaaaaa 22440
 ttagccaggg atggtggtgg gcacctgtaa tcccagctac tcaggaggct gaggtggcag 22500
 gatcacttga acccgggagg cgagggttgc agttagctga gatcgtgcca ctgtactcca 22560
 gcctaggtga cagagactct gtctcaaaaa aaaaaaaaaa ggcacgcacaa gataaaggat 22620
 cattagacac tagttagcct tcaattttcc tctttctct ctgtatccctt tcccattgaa tggaggtgaa 22680
 tcaagtccaa cccctacctg aactcttgat ctgtatccctt tcccattgaa tggaggtgaa 22740
 cttttgttcc tgtctcttct gtactgagtc tcttcctcta actcctgctt gtaatacgct 22800
 cagttatttc ttatcttcta aagtcaaact tctggacaaa aactccagtg tgctgttcaa 22860
 tactaaaaat agatttagaa gaaaaatatt ttccaaggtg aactgcacga taatgcgtca 22920
 gtagtgaagg gagcagccct ccagggggcg .tgcctgtcta .tctgttaacc acgttcatag 22980
 cagtagctg ctgtggtcag tgccatcaccc cttctcattt gatttcgtat gctctgtgag 23040
 gtagatagta ctttgacctc taaattatgt taccggaaata ttaaggttt atgtcattta 23100
 atattgaaca ataaagcaaa catagaatat tatgggatta gattgaagga agtaaaataa 23160
 taacataact tgctatacag tctccaacct attttcagt cgagcacata ctttcaacat 23220

s644PCT88.ST25

tttggaaataca	tttgtgcagt	aagaacttta	tgtttgata	ctattcaaaa	ttaagattta	23280
aacaaaaaat	ctgcatctta	ctgcatggct	tggccaattt	gccttactct	aacttacttt	23340
ataagcccat	aactttactg	atttttttt	caaatatttt	attatgaaaa	ttttactata	23400
ccacttagcc	tattacagtt	tatTTTgata	taatttgttt	agtacacttt	caaaaataat	23460
agttgacatc	tttctcatta	ataggtcaat	atgtgataaa	tgtttttaga	aaaggacgtt	23520
ttaaaaccaa	tgaataattc	agataacatt	ctttgtaaat	tatctaagcc	attctaaataa	23580
aattacctac	tttgaaagtt	aatttctaag	tataatgaat	atcagaggac	taaagataaa	23640
tgtatatgtg	tatatttata	tctagccata	tttgtgtcta	tgtatatata	catatataatg	23700
tatatcactc	tattttttt	tccactgttag	aaaaaaagctg	ctaaaagaga	agaagcatta	23760
gcatcatttgc	aggcctggaa	ggctatgaaa	gaaaaggaag	caaagaaaat	agctgccaaa	23820
aagaggcttg	aagaaaaaaa	caagaagaaa	actgaagaag	aaaatgctgc	aagaaaagga	23880
gaagcactac	aagtattcag	aactttgcac	atcttaatta	ttttaaaaca	tttgaatcc	23940
aaattaatga	ttaaccatat	tttttatttat	tttcaaataat	tcacagtaag	aaaattattc	24000
tgaacttttgc	caggcttttgc	aaaaatggaa	agagaaaaag	atggaatatc	ttaaagagaa	24060
aaatagaaaag	gagagagaat	atgaaagagc	aaagaaacag	aaagaggagg	aaactgttgc	24120
cgagaaaaag	aaagataatt	taactgctgt	tgagaaatgg	taatccaaaa	tcataaaatat	24180
tttgatataat	tttaaatttat	agtaacactt	caggattta	taaaatttat	ttacttgaaa	24240
tttagtaatgc	catttcaatt	tcattactgt	caaagatgta	ctaggaaatc	tttattatgt	24300
attttccttgc	aactctccag	tgttttatac	tatgctctat	aggaatgaaa	aaaaggaagc	24360
tttttcaag	caaaaggaaa	aagaaaaaat	aaatgagaaa	agaaaggaag	aactgaaaag	24420
agctgagaaa	aaagataaaag	ataaacaagc	tattaatgaa	tatgaaaaat	ggctggtagg	24480
tattatttgc	caatgcactt	tcgttttttgc	catgtacctt	ttgtgtcttt	tctgtcccta	24540
attctaattc	tatttgctcc	agacctactg	atcatttcta	cctggaaatct	gctttgttga	24600
attcaagctc	tcctcctgca	tatagcatat	tttcttgac	ttagtcattt	ctattaatgt	24660
ttctactatt	ccctcaaaaca	cccaggctga	aaacttgtta	taatcttctt	ccttacctgc	24720
atccccacat	ttaccattta	ctattcatgc	ccattttcc	tttgctgtga	ttctcacatc	24780
taacatagaa	agaagacaag	tttactatttgc	agggtactac	gtggtgaaac	ttggtcatga	24840
caaaaagtaa	cactgaactt	aatagtgaga	aaattattcc	atcttttatt	ctcttttgat	24900
gtttctgtat	acctcaagga	aatctctta	tttaggaatt	ttaatgaaa	gagagcaggt	24960
ttgagggttgc	ggaggagcaa	tagctagctg	aaccagatat	gtgtatataat	ttgatttcac	25020
tttacttatac	tttataaaag	ttacttttgc	ttgatgtcaa	gcaaaatatt	atttccatt	25080
tttagaatatc	aatataaata	tgcattttgt	ccatgtttat	ataagtaata	cattactatg	25140
aataaaatact	ttacataagt	aggtAACACA	ttcatatgaa	tagttAACAT	attcatatgaa	25200
ttcagcaacc	aaaattatag	tatTTTgca	ctagaagtct	atccagtcag	gtttcctatc	25260

s644PCT88.ST25

aaactttaaa acaactcata ccaatcaact aaatcatcca ggttggggatttgcatt 25320
 tctctggta gaattgagct tgaatatctt ttcatgtacaggccatttatctatta 25380
 ttttcctgt aaattgtcat ttcatagact ttgcacactt ttctattaga ttgttgggtt 25440
 ttttcctta ctggttctta gaatcttttg tttgtactg gggaaatttag cctatcattt 25500
 tttatatggg ttgcaaataat ttaccccccac tatattgttg gtttcccggc tttccttata 25560
 gtatctcatg ccatgaagaa tttaaatttt aggtgtcaga tttctgttttttttttgg 25620
 ctttgattt tcaagcatag ttgaaaagac ctacacaatt tgagattaa cagaattatc 25680
 ttatTTTct tctaacaact ttgtgacttt aatatcttaa tggtttaaca tttgttctgc 25740
 ttggaaTTTccctgataca tggggaaa tatgattca actttagttt ttccaaatgt 25800
 atcctttata aagtagccca ttttaccca ttgatttgag gtgctacttc tgttatatga 25860
 taccttctca tgTTTcggg tctgtttctt aactttctgt tccattggtc agtctcgta 25920
 ttccagtgcc acacttccat tattaggctt gatatgtcta aatatctgct tggattcatc 25980
 tccctttata gttttttttt cacagtcttt ctgaccagtc ttgtttttttt attttttcca 26040
 taaacttaag aatcagcagt agtttagaaag gtacatggga cccaaatggat cgatttaaag 26100
 ataggataaa aagataaaac aataataaac ttaagaaaca tgccagacca acataaagaa 26160
 aattgttagaa ctctcctgaa caacacaaat gaagacttga gaaaatggat cagaattgcc 26220
 catgcacaga aacacactta accttataat gatgttataa ggatgtcagc tctccctgaa 26280
 gtcatttaat gcaatcttaa caaaagccaa caggatttac tctgtgtgtt gagtttagta 26340
 ctgctatatg ctaattcgat gcagagaaat agtaataaaa taaggtatc aaaattgggt 26400
 caattttgaa tgaaaaaggt agtgtttcat gatgattcc ttaagttaat ctgttaata 26460
 atgctatgtt ctaaaaaaaaaa atttaaagtc cacttatatt aagaagatgt acactgactg 26520
 ctagtatcaa ttagggaaat taaatgtaaa catttgagtt ttccatttttta attccatatc 26580
 ttcatgaaaa tggaatagaa ttctttat aagtcacatt taggtataact gtttttaatt 26640
 atagcactta attacattgt catttttatac agtccctctga agaacaagaa ttccctcaaag 26700
 accaaagaca aaataacatg ttgatatct agtaaaatgt ctgcaaataat agtacaccta 26760
 taaacacata aacatacatg ttacagatcg gttctcccttc ttaccaaattt cttattgaaa 26820
 ttgtttgca gatagaatag aaaaattgcc cctgtatagg agtctaatga ctccagtttt 26880
 catggaaaac aacatctcaa gcttttata tacaaacttag tttgaacagt aagcattttgg 26940
 tgggttaattt ctttagggaa aagttatag ccaaagatca ggtttagacta aaatattttt 27000
 cttgccaattt accagattaa ttcatcatta ccttttagtaa gaaaataagc aaaaagctca 27060
 gttttccaca aataaatgtc tgaaggactt tttaacaagg ttcttttaat tactatcaag 27120
 gtgactattt attctttga actgatatta cagttatataattt aattgtctat ttgctaccct 27180
 ggctttacag ctccctgcta gtaagatgaa gcatattca agttactgcc ccctcatgtt 27240
 aagtgaaattt aaaaaagag atttattcag tcaatttctg tggacacagt ctggtcactg 27300

s644PCT88.ST25

cttttcttcc gccttagctag atggctgtc tctaaaatataaaatgatt gaagatgatc 27360
taattacagc ttgcctttc tcaattaaaa ttctgaaagg aagtttcctc ttgccttat 27420
tagaaatagc aagcaaacaa acatgcaagc attcttatga catggaatga ggatatgggt 27480
gttaacattg acaaaaaaca aacaaacctc ccacttcaact ttgtttgtta catgtgaatg 27540
gaaagcttgt cctgtattgc catattattc ttgtggcatt tatatatata ctgatgaaaa 27600
gatgcataca tacctaatac tttccataa tgcccttcct cccaagccat caacctgcag 27660
aggcaggtt cactaagggt ttcctgctc cttgaggaat atgagaaaaaa taccaagatg 27720
aagaaaccac caaaccttat agtggtagca gagacataaa gggacacctg gtgcctct 27780
tccatttctt gtctcctgcc ttctgccaag ccttagtcac aatggatatt ttgtttcct 27840
cccacagcac acatttttt tcccactctc agagccctca ccactactgt ttgcaagcaa 27900
agctctcccc cgatatttt cacgagtggc ttctcttatac catcatgtca cacttcaaag 27960
ggactttccc tgagtccatt ttttggaa agtaaatact ctttttatt ctttctcata 28020
gtttaaaac atgtttcaga gaaattcaca caatttgaa ttatctgtt tttttttct 28080
ttgtttctgt ccattttgaa agtccctgg gggacaggga ccatatctgt gtgttggat 28140
tttaaaaaat tattttatt tgcaaatgac acataaaaag tgcacatatt tatggaatac 28200
agtgtgatgt ttccatctac attgtataca ttgtgtaca atcagaaatg actcacaaag 28260
tagggaaaaa tggttgcata aaagatatac ttaatatttta ttataggaaa gtacacaaat 28320
tactaaaaat taaaggcaa taccatacat ttaaatggc caaataattg agcagaaaaat 28380
ttacaaaagg ctaaagaaat gttgaaaat gtgctcaagt tcaataataa agaaacatga 28440
ggcagaattt ttaactattt gtaaaaaatt tgaagtatct catactgtca tgacatattg 28500
aaactttgca cccagtaaac ttacttctga gaatttggc tcacgaagtc accaccaact 28560
tataacagtt actatatttgc agttataatt ataggctttt ttttcttattt tatacaattc 28620
tttttaatg ttttacttt taaagttaa aaaattaatg gatatttagta cttgcaattt 28680
gacaatgttt actaattttt ttcttggcatt catttttgc ttgtttgtt ttttggaca 28740
gggtctcaact ctgttgcata ggctggagtg cagtggtgca atctcgctc actgcaacct 28800
ccacctccccca ggctcaagca atcctcccat ctcagccctcc taagtaggtg ggactatagg 28860
catgcaccgc cacacctggc taattttgtt gttttttgtt agagatgtg tttcaccatg 28920
tttcccaggc tggctcgaa ctcccaggct caaacaatcc acccaccttta gtctcctaaa 28980
gttctggat tactggcatg agccaccatg cctggcccta cctgttattt ctttatgtatc 29040
tgttaaacta ggaagtgata tataaatatc ctataatggaa ttatattgtt cttcagcaag 29100
caacctgatt tgaaaataat aatcatatataat gtacataat ttatagtgtt ctatttctc 29160
tttaggaaaa taaggaaaaaa caagaaagaa ttgaacgaaa acagaagaaa cgtcattcct 29220
ttcttggaaag tgaggcactt cctccgtgga gccctccaag cagaactgtg ttcgc当地 29280
tgttttgata attcttagtttcc ttacatttttgc ttgttatttgc aatatttagcc 29340

s644PCT88.ST25

atagatttaa aaccattcaa ttattttatag ttagaggaat atattttaat taaatgccag 29400
 acactcctgc tgacaatgaa agaaatactt tggaaatgtaa tcagtgaaag cattttttg 29460
 aactgttagat aaactgcctc aaacaaagac ctaataatca gattgtttt accattaaga 29520
 tacataagat tttatcatgt cctgataatt cttatggtgg agtgattcat gatcttttc 29580
 attaagctct gtatgttatt taagtatatt taattccagt aataaaaagg aaatcatcta 29640
 ggtaccataa tgatagaaat tattcctttt gtggatgatt gtgaatctag attcaggaaa 29700
 tttaatgaag ggtcgctggg aagtgcgcat atattattcc ttctgaaact 29750

<210> 17

<211> 200

<212> DNA

<213> Homo sapiens

<400> 17
 acttccttcg tctgggtggt tgccccagcg acacgttggg ccgaagagcg gtgttggta 60
 cccgagagac ccggcggtgg ggaagtcact tcctcccgaa gacgctgttt cctagcaacc 120
 gccctccgcc tctgttatta gcccctcctc ctcgctcggt ccaggaccgg ctctgcgggc 180
 gccgccaggc ccagacccaag 200

<210> 18

<211> 139

<212> DNA

<213> Homo sapiens

<400> 18
 ctactatcg aagttgaatt ctaataatta gctattttat aaaggtaacg agaaaaaaata 60
 cactatgtct gatgaagttt ttagcaccac tttggcatat accaaagagtc caaaagttac 120
 caaaagaact actttccag 139

<210> 19

<211> 85

<212> DNA

<213> Homo sapiens

<400> 19
 gatgagctaa taagagcaat tacagctcgc tcagccagac aaaggagttc tgaatactca 60
 gatgactttg acagtgtatga gattg 85

S644PCT88.ST25

<210> 20
<211> 321
<212> DNA
<213> Homo sapiens

<400> 20
tttcctttagg tgattttct gacacttcag cagataaaa ttcagttaat aaaaaaatga 60
atgactttca tatatcagat gatgaagaaa agaatccttc aaaactattg ttttgaaaa 120
ccaataaaatc aaacggtaac ataaccaaag atgagccagt gtgtgccatc aaaaatgaag 180
aggaaatggc acctgatggg tgtgaagaca ttgttgtaaa atctttctct gaatctcaaa 240
ataaggatga ggaatttcaa aaagacaaaa taaaaaatgaa acctaaaccc agaattcttt 300
caattaaaaag cacatcttca g 321

<210> 21
<211> 227
<212> DNA
<213> Homo sapiens

<400> 21
cagaaaaacaa cagccttgac acagatgatc actttaaacc atcacctcg 60
tgcggccatc gaaaagtcac atggaggaga agatggact agaagataaa gaaactgccc 120
tcaggatgttgcac cttctccct tccaaacgccc aatggcatac 180
aattagaagc tgagaaaaaa gcattctctg aaaaccttga tcctgag 227

<210> 22
<211> 94
<212> DNA
<213> Homo sapiens

<400> 22
gattcatgct taacaagtct agcatcatca tcacttaaac aaattctgg agattcttt 60
tcaccaggat ctgagggaaa cgcatctgaa aag 94

<210> 23
<211> 248
<212> DNA

s644PCT88.ST25

<213> Homo sapiens

<400> 23
atccaaatga agaaatcact gaaaaccata attccttcaa atcagatgaa aataaagaga 60
attcattttc agcagaccat gtgactactg cagttgagaa atccaaggaa agtcaagtga 120
ctgctgatga ccttgaagaa gaaaaggcaa aagcggact gattatggat gatgacagaa 180
cagttgatcc actactatct aaatctcaga gtatcttaat atctaccagt gcaacagcat 240
cttcaaag 248

<210> 24

<211> 71

<212> DNA

<213> Homo sapiens

<400> 24
aaaacaattg aagatagaaaa tataaagaat aaaaagtcaa caaataatag agcatccagt 60
gcatctgccaa g 71

<210> 25

<211> 169

<212> DNA

<213> Homo sapiens

<400> 25
attaatgacc tctgagtttt tgaagaaatc tagttctaaa aggagaactc catcgacaac 60
tacctcttct cactattttag ggactttaaa agtcttgac caaaaacctt cacagaaaca 120
gagcatagaa cctgatagag cagataacat aagggcagct gtttatcag 169

<210> 26

<211> 90

<212> DNA

<213> Homo sapiens

<400> 26
gagtggtag aaaagaaaaa tgtgtattta catgaaatgc acagaataaa aagaattgaa 60
agtgaaaact taaggatcca aaatgaacag 90

<210> 27

s644PCT88.ST25

<211> 160

<212> DNA

<213> Homo sapiens

<400> 27
aaaaaaagctg ctaaaagaga agaaggcatta gcatcatttgcggaa ggctatgaaa 60
gaaaaggaag caaagaaaat agctgccaaa aagaggcttg aagaaaaaaaaa caagaagaaa 120
actgaagaag aaaatgctgc aagaaaagga gaagcactac 160

<210> 28

<211> 146

<212> DNA

<213> Homo sapiens

<400> 28
gcttttggaaa aatggaaaga gaaaaagatg gaatatctta aagagaaaaaa tagaaaggag 60
agagaatatg aaagagcaaa gaaacagaaa gaggagggaaa ctgttgccga gaaaaagaaaa 120
gataatttaa ctgctgttga gaaatg 146

<210> 29

<211> 133

<212> DNA

<213> Homo sapiens

<400> 29
gaatgaaaaa aaggaagctt ttttcaagca aaagggaaaaa gaaaaaataaa atgagaaaaag 60
aaaggaagaa ctgaaaagag ctgagaaaaaa agataaagat aaacaagcta ttaatgaata 120
tgaaaaatgg ctg 133

<210> 30

<211> 485

<212> DNA

<213> Homo sapiens

<400> 30
gaaaataagg aaaaacaaga aagaattgaa cgaaaacaga agaaacgtca ttccctttctt 60
gaaagtgagg cacttcctcc gtggagccct ccaagcagaa ctgtgttcgc aaaagtgttt 120
tgataattct agttcttaca ttatgggtt atttatcggt ttgccaatat tagccataga 180

s644PCT88.ST25

ttaaaaacca ttcaattatt tatagttaga ggaatataatt ttaattaaat gccagacact	240
cctgctgaca atgaaagaaa tactttggaa tgtaatcagt gaaagcattt ttttgaactg	300
tagataaaact gcctcaaaca aagacctaatt aatcagattt ttttaccat taagatacat	360
aagattttat catgtcctga taattcttatt ggtggagtga ttcatgatct ttttcattaa	420
gctctgtatg ttatthaagt atattaattt ccagtaataa aaaggaaatc atctaggtac	480
cataa	485

<210> 31

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 31 atgtctgatg aagtttttag cacc	24
--	----

<210> 32

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 32 aggcctcaaa tgatgctaat gc	22
--------------------------------------	----

<210> 33

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 33 atcatttgag gcctggaagg c	21
-------------------------------------	----

S644PCT88.ST25

<210> 34
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Amorce

<400> 34 aaacactttt gcgaacacag ttc 23

<210> 35
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Amorce

<400> 35 acaacgaaata acagagtgtc c 21

<210> 36
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Amorce

<400> 36 actcctgata aacagctgcc 20

<210> 37
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Amorce

s644PCT88.ST25

<400> 37
gccaccatgt ctgatgaagt ttttagcac 29

<210> 38

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 38
gaaacacttt tgcgAACACA gttc 24

<210> 39

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 39
taatgtctga tgaAGTTTT agcacc 26

<210> 40

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 40
tcaaaaACACT tttgcgaaca cagttc 26

<210> 41

<211> 25

<212> DNA

<213> Artificial sequence

s644PCT88.ST25

<220>

<223> Amorce

<400> 41

aatgtctgat gaagtttta gcacc

25

<210> 42

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 42

tcagcttgcc gtaggtggc

19

<210> 43

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> Amorce

<400> 43

atggtcctgc tggagttcg

19

<210> 44

<211> 391

<212> DNA

<213> Murinae gen. sp.

<400> 44

aaagaagtga agacagaaaac acgaagaata aaaagacaac gaataacaga gtgtccagtg

60

cctctggcag gctgatgacc tctgagttt taaagagatc cggccccaca aaaagaagtc

120

catctgcagc tacctcctca cactattag ggagttgaa agtcttggac cagaagcaac

180

cacgaaagca gagccttagag ccagacaagg ctgatcacat aagggcagct gtttatcagg

240

agtggtaga aaagaaaaat gtgtattac atgaaatgca cagaataaaa agaattgaaa

300

gcgaaaacctt gaggatccaa aatgaacaga aaaaagctgc taagagagag gaagccctgg

360

catcatggaa ggcctggaa gcaatgaaag a

391

s644PCT88.ST25

<210> 45
<211> 2767
<212> DNA
<213> *Mus musculus*

<220>
<221> CDS
<222> (204)..(2147)
<223>

s644PCT88.ST25

gtc aaa aga agc acg tcc tcg ggg gaa acc agc agc ggt ctt gat gca Val Lys Arg Ser Thr Ser Ser Gly Glu Thr Ser Ser Gly Leu Asp Ala 155 160 165 170	713
gat ggc cac ttt aag cct tca ccc cag cca agg agc atg tta aaa aag Asp Gly His Phe Lys Pro Ser Pro Gln Pro Arg Ser Met Leu Lys Lys 175 180 185	761
agc agc cac act gag gag gga gtc aga cca gga gtt gat aaa gaa cat Ser Ser His Thr Glu Glu Gly Val Arg Pro Gly Val Asp Lys Glu His 190 195 200	809
tcc ata agc gaa gcc tct gct ccc aca cct tcc ctt cca agg cag aat Ser Ile Ser Glu Ala Ser Ala Pro Thr Pro Ser Leu Pro Arg Gln Asn 205 210 215	857
ggc aca gag ttg caa act gag gaa aaa ata tac tcg gaa aac ctc gat Gly Thr Glu Leu Gln Thr Glu Glu Lys Ile Tyr Ser Glu Asn Leu Asp 220 225 230	905
ctt gag gac tca ctc tta caa agt ctg acc tca tct tcc ttc aaa gaa Leu Glu Asp Ser Leu Leu Gln Ser Leu Thr Ser Ser Ser Phe Lys Glu 235 240 245 250	953
agc ccc gga ggt tgc aca tca cca gga tct cag gaa aag gtg ccc ata Ser Pro Gly Cys Thr Ser Pro Gly Ser Gln Glu Lys Val Pro Ile 255 260 265	1001
aaa gat cat gat gga gaa cct act gaa atc tgg gat tcc ttg cta tca Lys Asp His Asp Gly Glu Pro Thr Glu Ile Trp Asp Ser Leu Leu Ser 270 275 280	1049
aat gaa aat gaa gga agt tct gtt ttg gtg aac tgt gtt act cct gaa Asn Glu Asn Glu Gly Ser Ser Val Leu Val Asn Cys Val Thr Pro Glu 285 290 295	1097
ctc gag cag ccc aag gac ggt cag gtg gca gct gac gac ctt gag gaa Leu Glu Gln Pro Lys Asp Gly Gln Val Ala Ala Asp Asp Leu Glu Glu 300 305 310	1145
gaa aga gag aag ggt gga ttt aca gaa gat gac ctc acc act gac ccg Glu Arg Glu Lys Gly Gly Phe Thr Glu Asp Asp Leu Thr Thr Asp Pro 315 320 325 330	1193
ctg ctc tcc acg tcc ccg agt gtc ata aca ccc act gag cca gca gag Leu Leu Ser Thr Ser Pro Ser Val Ile Thr Pro Thr Glu Pro Ala Glu 335 340 345	1241
ccg gcc aag aaa gca aat gaa gac aga aac acg aag aat aaa aag aca Pro Ala Lys Lys Ala Asn Glu Asp Arg Asn Thr Lys Asn Lys Lys Thr 350 355 360	1289
acg aat aac aga gtg tcc agt gcc tct ggc agc agg ctg atg acc tct Thr Asn Asn Arg Val Ser Ser Ala Ser Gly Ser Arg Leu Met Thr Ser 365 370 375	1337
gag ttt tta aag aga tcc ggt ccc aca aaa aga agt cca tct gca gct Glu Phe Leu Lys Arg Ser Gly Pro Thr Lys Arg Ser Pro Ser Ala Ala 380 385 390	1385
acc tcc tca cac tat tta ggg agt ttg aaa gtc ttg gac cag aag caa Thr Ser Ser His Tyr Leu Gly Ser Leu Lys Val Leu Asp Gln Lys Gln 395 400 405 410	1433
cca cgg aag cag agc cta gag cca gac aag gct gat cac ata agg gca Pro Arg Lys Gln Ser Leu Glu Pro Asp Lys Ala Asp His Ile Arg Ala 415 420 425	1481

s644PCT88.ST25

gct gtt tat cag gag tgg tta gaa aag aaa aat gtg tat tta cat gaa Ala Val Tyr Gln Glu Trp Leu Glu Lys Lys Asn Val Tyr Leu His Glu 430 435 440	1529
atg cac aga ata aaa aga att gaa agc gaa aac ttg agg atc caa aat Met His Arg Ile Lys Arg Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn 445 450 455	1577
gaa cag aaa aaa gct gct aag aga gag gaa gcc ctg gca tca ttt gag Glu Gln Lys Lys Ala Ala Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu 460 465 470	1625
gcc tgg aag gca atg aaa gag aag gaa gca aag aga ata gct gca aaa Ala Trp Lys Ala Met Lys Glu Lys Glu Ala Lys Arg Ile Ala Ala Lys 475 480 485 490	1673
aag agg ctg gag gaa aag aac aag aag aaa aca gaa gaa gaa aat gcc Lys Arg Leu Glu Glu Lys Asn Lys Lys Lys Thr Glu Glu Glu Asn Ala 495 500 505	1721
atg agg aaa ggc gag gcc ctg caa gca ttt gaa aaa tgg aaa gag aaa Met Arg Lys Gly Glu Ala Leu Gln Ala Phe Glu Lys Trp Lys Glu Lys 510 515 520	1769
aag cta gaa tac ctc aaa gag aag acc agg agg gag aaa gaa tat gaa Lys Leu Glu Tyr Leu Lys Glu Lys Thr Arg Arg Glu Lys Glu Tyr Glu 525 530 535	1817
aga gca aag aaa cag aaa gaa gag gaa gcg gtt gct gag aaa aag aaa Arg Ala Lys Lys Gln Lys Glu Glu Glu Ala Val Ala Glu Lys Lys Lys 540 545 550	1865
gac agt tta act gct ttt gaa aaa tgg agt gag aga aag gaa gct ctc Asp Ser Leu Thr Ala Phe Glu Lys Trp Ser Glu Arg Lys Glu Ala Leu 555 560 565 570	1913
ctc aag caa aag gag aag gag aaa ata aat gag aga aga aag gaa gag Leu Lys Gln Lys Glu Lys Glu Lys Ile Asn Glu Arg Arg Lys Glu Glu 575 580 585	1961
ctg aag aga gcc gag aag aaa gac aaa gac aag caa gcc atc agt gaa Leu Lys Arg Ala Glu Lys Lys Asp Lys Asp Lys Gln Ala Ile Ser Glu 590 595 600	2009
tac gaa aag tgg ctg gaa aag aaa gaa agg caa gaa aga att gaa cgg Tyr Glu Lys Trp Leu Glu Lys Lys Glu Arg Gln Glu Arg Ile Glu Arg 605 610 615	2057
aaa cag aag aag cgc cac tcc ttc ctt gag agc gag aca cac cca cca Lys Gln Lys Lys Arg His Ser Phe Leu Glu Ser Glu Thr His Pro Pro 620 625 630	2105
tgg agt cct ccg agc aga act gcg ccc tca aaa gta ttt tga Trp Ser Pro Pro Ser Arg Thr Ala Pro Ser Lys Val Phe 635 640 645	2147
tgtttctggc tcttgatttt tttttcagtt caccaactgt actcatggat ttaaaaacgag	2207
tcatctcatt atttgtggtt agaagactct atgtcacttc cctgcaggag cttctgtgga	2267
gcatgaaaga gatactttgc agttaatca gtggaaacat tttctgaagt gtcctcatca	2327
gtttgctggc acaatccaga cgcatgaagc tttattatga cctgaacagt ctgggtgtggc	2387
gtgattcgtg gtcactgtcg ctgagttcgg agtctttta aagaatgttt gatcccacta	2447
atgaaagaat gccagctaga taccacaatc gtagagatga ctcggctgt ggaagtctgt	2507

s644PCT88.ST25

gcttcttagag	tgtatggggcagg	ccatggggcat	gttatcttt	2567
ctaactccag	ttcttcaggt	cacagaagta	tcttgctgt	2627
gttgaggcca	cagaactcta	gtcagtcact	ttagtaaaga	2687
ctcggtgtgg	tttgcccttct	tgaggcttac	ctgacaatcg	2747
tcacttctgg	aatgttcttt			2767

<210> 46

<211> 647

<212> PRT

<213> Mus musculus

<400> 46

Met	Ser	Asp	Glu	Ile	Phe	Ser	Thr	Thr	Leu	Ala	Tyr	Thr	Lys	Ser	Pro
1				5					10				15		

Lys	Ala	Thr	Lys	Arg	Thr	Ser	Phe	Gln	Asp	Glu	Leu	Ile	Arg	Ala	Ile
			20			25						30			

Thr	Ala	Arg	Ser	Ala	Arg	Gln	Arg	Ser	Ser	Glu	Tyr	Ser	Asp	Asp	Phe
35					40					45					

Asp	Ser	Asp	Glu	Ile	Val	Ser	Leu	Gly	Glu	Phe	Ser	Asp	Thr	Ser	Thr
50					55					60					

Asp	Glu	Ser	Leu	Val	Arg	Lys	Lys	Met	Asn	Asp	Phe	His	Ile	Ser	Asp
65					70			75			80				

Asp	Glu	Glu	Lys	Asn	Ser	Pro	Arg	Leu	Ser	Phe	Leu	Lys	Thr	Lys	Lys
85						90						95			

Val	Asn	Arg	Ala	Ile	Ser	Asn	Asp	Ala	Leu	Asp	Ser	Ser	Thr	Pro	Gly
100						105					110				

Ser	Glu	Gly	Ser	Ser	Pro	Asp	Ala	Gln	Glu	Asp	Val	Thr	Gly	Asp	Ser
115						120					125				

Leu	Pro	Lys	Ser	Gln	Asn	Asp	Asp	Arg	Glu	Val	Gly	Arg	Glu	Ile	Ile
130					135					140					

Thr	Val	Lys	Pro	Thr	Pro	Arg	Met	His	Pro	Val	Lys	Arg	Ser	Thr	Ser
145					150			155			160				

Ser	Gly	Glu	Thr	Ser	Ser	Gly	Leu	Asp	Ala	Asp	Gly	His	Phe	Lys	Pro
165						170					175				

s644PCT88.ST25

Ser Pro Gln Pro Arg Ser Met Leu Lys Lys Ser Ser His Thr Glu Glu
 180 185 190

Gly Val Arg Pro Gly Val Asp Lys Glu His Ser Ile Ser Glu Ala Ser
 195 200 205

Ala Pro Thr Pro Ser Leu Pro Arg Gln Asn Gly Thr Glu Leu Gln Thr
 210 215 220

Glu Glu Lys Ile Tyr Ser Glu Asn Leu Asp Leu Glu Asp Ser Leu Leu
 225 230 235 240

Gln Ser Leu Thr Ser Ser Ser Phe Lys Glu Ser Pro Gly Gly Cys Thr
 245 250 255

Ser Pro Gly Ser Gln Glu Lys Val Pro Ile Lys Asp His Asp Gly Glu
 260 265 270

Pro Thr Glu Ile Trp Asp Ser Leu Leu Ser Asn Glu Asn Glu Gly Ser
 275 280 285

Ser Val Leu Val Asn Cys Val Thr Pro Glu Leu Glu Gln Pro Lys Asp
 290 295 300

Gly Gln Val Ala Ala Asp Asp Leu Glu Glu Glu Arg Glu Lys Gly Gly
 305 310 315 320

Phe Thr Glu Asp Asp Leu Thr Thr Asp Pro Leu Leu Ser Thr Ser Pro
 325 330 335

Ser Val Ile Thr Pro Thr Glu Pro Ala Glu Pro Ala Lys Lys Ala Asn
 340 345 350

Glu Asp Arg Asn Thr Lys Asn Lys Lys Thr Thr Asn Asn Arg Val Ser
 355 360 365

Ser Ala Ser Gly Ser Arg Leu Met Thr Ser Glu Phe Leu Lys Arg Ser
 370 375 380

Gly Pro Thr Lys Arg Ser Pro Ser Ala Ala Thr Ser Ser His Tyr Leu
 385 390 395 400

Gly Ser Leu Lys Val Leu Asp Gln Lys Gln Pro Arg Lys Gln Ser Leu
 405 410 415

Glu Pro Asp Lys Ala Asp His Ile Arg Ala Ala Val Tyr Gln Glu Trp
 420 425 430

Leu Glu Lys Lys Asn Val Tyr Leu His Glu Met His Arg Ile Lys Arg
 435 440 445

s644PCT88.ST25

Ile	Glu	Ser	Glu	Asn	Leu	Arg	Ile	Gln	Asn	Glu	Gln	Lys	Lys	Ala	Ala
450					455							460			

Lys	Arg	Glu	Glu	Ala	Leu	Ala	Ser	Phe	Glu	Ala	Trp	Lys	Ala	Met	Lys
465					470							475			480

Glu	Lys	Glu	Ala	Lys	Arg	Ile	Ala	Ala	Lys	Lys	Arg	Leu	Glu	Glu	Lys
						485			490						495

Asn	Lys	Lys	Lys	Thr	Glu	Glu	Asn	Ala	Met	Arg	Lys	Gly	Glu	Ala	
							500		505						

Leu	Gln	Ala	Phe	Glu	Lys	Trp	Lys	Glu	Lys	Lys	Leu	Glu	Tyr	Leu	Lys
						515		520							525

Glu	Lys	Thr	Arg	Arg	Glu	Lys	Glu	Tyr	Glu	Arg	Ala	Lys	Lys	Gln	Lys
						530		535							540

Glu	Glu	Glu	Ala	Val	Ala	Glu	Lys	Lys	Asp	Ser	Leu	Thr	Ala	Phe	
						545		550							560

Glu	Lys	Trp	Ser	Glu	Arg	Lys	Glu	Ala	Leu	Leu	Lys	Gln	Lys	Glu	Lys
						565		570							575

Glu	Lys	Ile	Asn	Glu	Arg	Arg	Lys	Glu	Glu	Leu	Lys	Arg	Ala	Glu	Lys
							580		585						590

Lys	Asp	Lys	Asp	Lys	Gln	Ala	Ile	Ser	Glu	Tyr	Glu	Lys	Trp	Leu	Glu
							595		600						

Lys	Lys	Glu	Arg	Gln	Glu	Arg	Ile	Glu	Arg	Lys	Gln	Lys	Lys	Arg	His
							610		615						620

Ser	Phe	Leu	Glu	Ser	Glu	Thr	His	Pro	Pro	Trp	Ser	Pro	Pro	Ser	Arg
						625		630							640

Thr	Ala	Pro	Ser	Lys	Val	Phe									
						645									

<210> 47

<211> 647

<212> PRT

<213> artificial sequence

<400> 47

Met	Ser	Asp	Glu	Ile	Phe	Ser	Thr	Thr	Leu	Ala	Tyr	Thr	Lys	Ser	Pro
1									5						15

S644PCT88.ST25

Lys Ala Thr Lys Arg Thr Ser Phe Gln Asp Glu Leu Ile Arg Ala Ile
20 25 30

Thr Ala Arg Ser Ala Arg Gln Arg Ser Ser Glu Tyr Ser Asp Asp Phe
35 40 45

Asp Ser Asp Glu Ile Val Ser Leu Gly Glu Phe Ser Asp Thr Ser Thr
50 55 60

Asp Glu Ser Leu Val Arg Lys Lys Met Asn Asp Phe His Ile Ser Asp
65 70 75 80

Asp Glu Glu Lys Asn Ser Pro Arg Leu Ser Phe Leu Lys Thr Lys Lys
85 90 95

Val Asn Arg Ala Ile Ser Asn Asp Ala Leu Asp Ser Ser Thr Pro Gly
100 105 110

Ser Glu Gly Ser Ser Pro Asp Ala Gln Glu Asp Val Thr Gly Asp Ser
115 120 125

Leu Pro Lys Ser Gln Asn Asp Asp Arg Glu Val Gly Arg Glu Ile Ile
130 135 140

Thr Val Lys Pro Thr Pro Arg Met His Pro Val Lys Arg Ser Thr Ser
145 150 155 160

Ser Gly Glu Thr Ser Ser Gly Leu Asp Ala Asp Gly His Phe Lys Pro
165 170 175

Ser Pro Gln Pro Arg Ser Met Leu Lys Lys Ser Ser His Thr Glu Glu
180 185 190

Gly Val Arg Pro Gly Val Asp Lys Glu His Ser Ile Ser Glu Ala Ser
195 200 205

Ala Pro Thr Pro Ser Leu Pro Arg Gln Asn Gly Thr Glu Leu Gln Thr
210 215 220

Glu Glu Lys Ile Tyr Ser Glu Asn Leu Asp Leu Glu Asp Ser Leu Leu
225 230 235 240

Gln Ser Leu Thr Ser Ser Ser Phe Lys Glu Ser Pro Gly Gly Cys Thr
245 250 255

Ser Pro Gly Ser Gln Glu Lys Val Pro Ile Lys Asp His Asp Gly Glu
260 265 270

Pro Thr Glu Ile Trp Asp Ser Leu Leu Ser Asn Glu Asn Glu Gly Ser
275 280 285

s644PCT88.ST25

Ser Val Leu Val Asn Cys Val Thr Pro Glu Leu Glu Gln Pro Lys Asp
 290 295 300

Gly Gln Val Ala Ala Asp Asp Leu Glu Glu Glu Arg Glu Lys Gly Gly
 305 310 315 320

Phe Thr Glu Asp Asp Leu Thr Thr Asp Pro Leu Leu Ser Thr Ser Pro
 325 330 335

Ser Val Ile Thr Pro Thr Glu Pro Ala Glu Pro Ala Lys Lys Ala Asn
 340 345 350

Glu Asp Arg Asn Thr Lys Asn Lys Lys Thr Thr Asn Asn Arg Val Ser
 355 360 365

Ser Ala Ser Gly Ser Arg Leu Met Thr Ser Glu Phe Leu Lys Arg Ser
 370 375 380

Gly Pro Thr Lys Arg Ser Pro Ser Ala Ala Thr Ser Ser His Tyr Leu
 385 390 395 400

Gly Ser Leu Lys Val Leu Asp Gln Lys Gln Pro Arg Lys Gln Ser Leu
 405 410 415

Glu Pro Asp Lys Ala Asp His Ile Arg Ala Ala Val Tyr Gln Glu Trp
 420 425 430

Leu Glu Lys Lys Asn Val Tyr Leu His Glu Met His Arg Ile Lys Arg
 435 440 445

Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn Glu Gln Lys Lys Ala Ala
 450 455 460

Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu Ala Trp Lys Ala Met Lys
 465 470 475 480

Glu Lys Glu Ala Lys Arg Ile Ala Ala Lys Lys Arg Leu Glu Glu Lys
 485 490 495

Asn Lys Lys Lys Thr Glu Glu Asn Ala Met Arg Lys Gly Glu Ala
 500 505 510

Leu Gln Ala Phe Glu Lys Trp Lys Glu Lys Lys Leu Glu Tyr Leu Lys
 515 520 525

Glu Lys Thr Arg Arg Glu Lys Glu Tyr Glu Arg Ala Lys Lys Gln Lys
 530 535 540

Glu Glu Glu Ala Val Ala Glu Lys Lys Lys Asp Ser Leu Thr Ala Phe
 545 550 555 560

Glu Lys Trp Ser Glu Arg Lys Glu Ala Leu Leu Lys Gln Lys Glu Lys
565 570 575

Glu Lys Ile Asn Glu Arg Arg Lys Glu Glu Leu Lys Arg Ala Glu Lys
580 585 590

Lys Asp Lys Asp Lys Gln Ala Ile Ser Glu Tyr Glu Lys Trp Leu Glu
595 600 605

Lys Lys Glu Arg Gln Glu Arg Ile Glu Arg Lys Gln Lys Lys Arg His
610 615 620

Ser Phe Leu Glu Ser Glu Thr His Pro Pro Trp Ser Pro Pro Ser Arg
625 630 635 640

Thr Ala Pro Ser Lys Val Phe
645

<210> 48

<211> 344

<212> PRT

<213> artificial sequence

<220>

<223> hASAP peptide mutant 411-647

<400> 48

Glu Ser Gln Val Thr Ala Asp Asp Leu Glu Glu Glu Lys Ala Lys Ala
1 5 10 15

Glu Leu Ile Met Asp Asp Asp Arg Thr Val Asp Pro Leu Leu Ser Lys
20 25 30

Ser Gln Ser Ile Leu Ile Ser Thr Ser Ala Thr Ala Ser Ser Lys Lys
35 40 45

Thr Ile Glu Asp Arg Asn Ile Lys Asn Lys Ser Thr Asn Asn Arg
50 55 60

Ala Ser Ser Ala Ser Ala Arg Leu Met Thr Ser Glu Phe Leu Lys Lys
65 70 75 80

Ser Ser Ser Lys Arg Arg Thr Pro Ser Thr Thr Ser Ser His Tyr
85 90 95

Leu Gly Thr Leu Lys Val Leu Asp Gln Lys Pro Ser Gln Lys Gln Ser
100 105 110

s644PCT88.ST25

Ile	Glu	Pro	Asp	Arg	Ala	Asp	Asn	Ile	Arg	Ala	Ala	Val	Tyr	Gln	Glu
115								120						125	
Trp	Leu	Glu	Lys	Lys	Asn	Val	Tyr	Leu	His	Glu	Met	His	Arg	Ile	Lys
130						135				140					
Arg	Ile	Glu	Ser	Glu	Asn	Leu	Arg	Ile	Gln	Asn	Glu	Gln	Lys	Lys	Ala
145						150			155						160
Ala	Lys	Arg	Glu	Glu	Ala	Leu	Ala	Ser	Phe	Glu	Ala	Trp	Lys	Ala	Met
								165		170					175
Lys	Glu	Lys	Glu	Ala	Lys	Lys	Ile	Ala	Ala	Lys	Lys	Arg	Leu	Glu	Glu
							180		185				190		
Lys	Asn	Lys	Lys	Thr	Glu	Glu	Glu	Asn	Ala	Ala	Arg	Lys	Gly	Glu	
						195		200			205				
Ala	Leu	Gln	Ala	Phe	Glu	Lys	Trp	Lys	Glu	Lys	Lys	Met	Glu	Tyr	Leu
						210		215				220			
Lys	Glu	Lys	Asn	Arg	Lys	Glu	Arg	Glu	Tyr	Glu	Arg	Ala	Lys	Lys	Gln
						225		230			235				240
Lys	Glu	Glu	Glu	Thr	Val	Ala	Glu	Lys	Lys	Lys	Asp	Asn	Leu	Thr	Ala
						245		250					255		
Val	Glu	Lys	Trp	Asn	Glu	Lys	Lys	Glu	Ala	Phe	Phe	Lys	Gln	Lys	Lys
						260		265				270			
Lys	Glu	Lys	Ile	Asn	Glu	Lys	Arg	Lys	Glu	Glu	Leu	Lys	Arg	Ala	Glu
						275		280				285			
Lys	Lys	Asp	Lys	Asp	Lys	Gln	Ala	Ile	Asn	Glu	Tyr	Glu	Lys	Trp	Leu
						290		295			300				
Glu	Asn	Lys	Glu	Lys	Gln	Glu	Arg	Ile	Glu	Arg	Lys	Gln	Lys	Lys	Arg
						305		310			315				320
His	Ser	Phe	Leu	Glu	Ser	Glu	Ala	Leu	Pro	Pro	Trp	Ser	Pro	Pro	Ser
								325		330					335
Arg	Thr	Val	Phe	Ala	Lys	Val	Phe								
						340									

<210> 49
<211> 237
<212> PRT
<213> artificial sequence

s644PCT88.ST25

<220>

<223> hASAP mutant 411-647

<400> 49

Ser Gln Lys Gln Ser Ile Glu Pro Asp Arg Ala Asp Asn Ile Arg Ala
1 5 10 15

Ala Val Tyr Gln Glu Trp Leu Glu Lys Lys Asn Val Tyr Leu His Glu
20 25 30

Met His Arg Ile Lys Arg Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn
35 40 45

Glu Gln Lys Lys Ala Ala Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu
50 55 60

Ala Trp Lys Ala Met Lys Glu Lys Glu Ala Lys Lys Ile Ala Ala Lys
65 70 75 80

Lys Arg Leu Glu Glu Lys Asn Lys Lys Thr Glu Glu Glu Asn Ala
85 90 95

Ala Arg Lys Gly Glu Ala Leu Gln Ala Phe Glu Lys Trp Lys Glu Lys
100 105 110

Lys Met Glu Tyr Leu Lys Glu Lys Asn Arg Lys Glu Arg Glu Tyr Glu
115 120 125

Arg Ala Lys Lys Gln Lys Glu Glu Glu Thr Val Ala Glu Lys Lys Lys
130 135 140

Asp Asn Leu Thr Ala Val Glu Lys Trp Asn Glu Lys Lys Glu Ala Phe
145 150 155 160

Phe Lys Gln Lys Lys Lys Glu Lys Ile Asn Glu Lys Arg Lys Glu Glu
165 170 175

Leu Lys Arg Ala Glu Lys Lys Asp Lys Asp Lys Gln Ala Ile Asn Glu
180 185 190

Tyr Glu Lys Trp Leu Glu Asn Lys Glu Lys Gln Glu Arg Ile Glu Arg
195 200 205

Lys Gln Lys Lys Arg His Ser Phe Leu Glu Ser Glu Ala Leu Pro Pro
210 215 220

Trp Ser Pro Pro Ser Arg Thr Val Phe Ala Lys Val Phe
225 230 235

s644PCT88.ST25

<210> 50

<211> 170

<212> PRT

<213> artificial sequence

<220>

<223> hASAP mutant 478-647

<400> 50

Ala Met Lys Glu Lys Glu Ala Lys Lys Ile Ala Ala Lys Lys Arg Leu
1 5 10 15

Glu Glu Lys Asn Lys Lys Thr Glu Glu Glu Asn Ala Ala Arg Lys
20 25 30

Gly Glu Ala Leu Gln Ala Phe Glu Lys Trp Lys Glu Lys Lys Met Glu
35 40 45

Tyr Leu Lys Glu Lys Asn Arg Lys Glu Arg Glu Tyr Glu Arg Ala Lys
50 55 60

Lys Gln Lys Glu Glu Glu Thr Val Ala Glu Lys Lys Lys Asp Asn Leu
65 70 75 80

Thr Ala Val Glu Lys Trp Asn Glu Lys Lys Glu Ala Phe Phe Lys Gln
85 90 95

Lys Lys Lys Glu Lys Ile Asn Glu Lys Arg Lys Glu Glu Leu Lys Arg
100 105 110

Ala Glu Lys Lys Asp Asp Lys Gln Ala Ile Asn Glu Tyr Glu Lys
115 120 125

Trp Leu Glu Asn Lys Glu Lys Gln Glu Arg Ile Glu Arg Lys Gln Lys
130 135 140

Lys Arg His Ser Phe Leu Glu Ser Glu Ala Leu Pro Pro Trp Ser Pro
145 150 155 160

Pro Ser Arg Thr Val Phe Ala Lys Val Phe
165 170

<210> 51

<211> 477

<212> PRT

<213> artificial sequence

s644PCT88.ST25

<220>

<223> hASAP mutant 1-477

<400> 51

Met Ser Asp Glu Val Phe Ser Thr Thr Leu Ala Tyr Thr Lys Ser Pro
1 5 10 15

Lys Val Thr Lys Arg Thr Thr Phe Gln Asp Glu Leu Ile Arg Ala Ile
20 25 30

Thr Ala Arg Ser Ala Arg Gln Arg Ser Ser Glu Tyr Ser Asp Asp Phe
35 40 45

Asp Ser Asp Glu Ile Val Ser Leu Gly Asp Phe Ser Asp Thr Ser Ala
50 55 60

Asp Glu Asn Ser Val Asn Lys Lys Met Asn Asp Phe His Ile Ser Asp
65 70 75 80

Asp Glu Glu Lys Asn Pro Ser Lys Leu Leu Phe Leu Lys Thr Asn Lys
85 90 95

Ser Asn Gly Asn Ile Thr Lys Asp Glu Pro Val Cys Ala Ile Lys Asn
100 105 110

Glu Glu Glu Met Ala Pro Asp Gly Cys Glu Asp Ile Val Val Lys Ser
115 120 125

Phe Ser Glu Ser Gln Asn Lys Asp Glu Glu Phe Glu Lys Asp Lys Ile
130 135 140

Lys Met Lys Pro Lys Pro Arg Ile Leu Ser Ile Lys Ser Thr Ser Ser
145 150 155 160

Ala Glu Asn Asn Ser Leu Asp Thr Asp Asp His Phe Lys Pro Ser Pro
165 170 175

Trp Pro Arg Ser Met Leu Lys Lys Ser His Met Glu Glu Lys Asp
180 185 190

Gly Leu Glu Asp Lys Glu Thr Ala Leu Ser Glu Glu Leu Glu Leu His
195 200 205

Ser Ala Pro Ser Ser Leu Pro Thr Pro Asn Gly Ile Gln Leu Glu Ala
210 215 220

Glu Lys Lys Ala Phe Ser Glu Asn Leu Asp Pro Glu Asp Ser Cys Leu
225 230 235 240

S644PCT88.ST25

Thr Ser Leu Ala Ser Ser Ser Leu Lys Gln Ile Leu Gly Asp Ser Phe
245 250 255

Ser Pro Gly Ser Glu Gly Asn Ala Ser Gly Lys Asp Pro Asn Glu Glu
260 265 270

Ile Thr Glu Asn His Asn Ser Leu Lys Ser Asp Glu Asn Lys Glu Asn
275 280 285

Ser Phe Ser Ala Asp His Val Thr Thr Ala Val Glu Lys Ser Lys Glu
290 295 300

Ser Gln Val Thr Ala Asp Asp Leu Glu Glu Lys Ala Lys Ala Glu
305 310 315 320

Leu Ile Met Asp Asp Asp Arg Thr Val Asp Pro Leu Leu Ser Lys Ser
325 330 335

Gln Ser Ile Leu Ile Ser Thr Ser Ala Thr Ala Ser Ser Lys Lys Thr
340 345 350

Ile Glu Asp Arg Asn Ile Lys Asn Lys Ser Thr Asn Asn Arg Ala
355 360 365

Ser Ser Ala Ser Ala Arg Leu Met Thr Ser Glu Phe Leu Lys Lys Ser
370 375 380

Ser Ser Lys Arg Arg Thr Pro Ser Thr Thr Ser Ser His Tyr Leu
385 390 395 400

Gly Thr Leu Lys Val Leu Asp Gln Lys Pro Ser Gln Lys Gln Ser Ile
405 410 415

Glu Pro Asp Arg Ala Asp Asn Ile Arg Ala Ala Val Tyr Gln Glu Trp
420 425 430

Leu Glu Lys Lys Asn Val Tyr Leu His Glu Met His Arg Ile Lys Arg
435 440 445

Ile Glu Ser Glu Asn Leu Arg Ile Gln Asn Glu Gln Lys Lys Ala Ala
450 455 460

Lys Arg Glu Glu Ala Leu Ala Ser Phe Glu Ala Trp Lys
465 470 475

<210> 52

<211> 418

<212> PRT

<213> artificial sequence

S644PCT88.ST25

<220>

<223> hASAP mutant 1-418

<400> 52

Met Ser Asp Glu Val Phe Ser Thr Thr Leu Ala Tyr Thr Lys Ser Pro
1 5 10 15

Lys Val Thr Lys Arg Thr Thr Phe Gln Asp Glu Leu Ile Arg Ala Ile
20 25 30

Thr Ala Arg Ser Ala Arg Gln Arg Ser Ser Glu Tyr Ser Asp Asp Phe
35 40 45

Asp Ser Asp Glu Ile Val Ser Leu Gly Asp Phe Ser Asp Thr Ser Ala
50 55 60

Asp Glu Asn Ser Val Asn Lys Lys Met Asn Asp Phe His Ile Ser Asp
65 70 75 80

Asp Glu Glu Lys Asn Pro Ser Lys Leu Leu Phe Leu Lys Thr Asn Lys
85 90 95

Ser Asn Gly Asn Ile Thr Lys Asp Glu Pro Val Cys Ala Ile Lys Asn
100 105 110

Glu Glu Glu Met Ala Pro Asp Gly Cys Glu Asp Ile Val Val Lys Ser
115 120 125

Phe Ser Glu Ser Gln Asn Lys Asp Glu Glu Phe Glu Lys Asp Lys Ile
130 135 140

Lys Met Lys Pro Lys Pro Arg Ile Leu Ser Ile Lys Ser Thr Ser Ser
145 150 155 160

Ala Glu Asn Asn Ser Leu Asp Thr Asp Asp His Phe Lys Pro Ser Pro
165 170 175

Trp Pro Arg Ser Met Leu Lys Lys Ser His Met Glu Glu Lys Asp
180 185 190

Gly Leu Glu Asp Lys Glu Thr Ala Leu Ser Glu Glu Leu Glu Leu His
195 200 205

Ser Ala Pro Ser Ser Leu Pro Thr Pro Asn Gly Ile Gln Leu Glu Ala
210 215 220

Glu Lys Lys Ala Phe Ser Glu Asn Leu Asp Pro Glu Asp Ser Cys Leu
225 230 235 240

S644PCT88.ST25

Thr Ser Leu Ala Ser Ser Ser Leu Lys Gln Ile Leu Gly Asp Ser Phe
 245 250 255

Ser Pro Gly Ser Glu Gly Asn Ala Ser Gly Lys Asp Pro Asn Glu Glu
 260 265 270

Ile Thr Glu Asn His Asn Ser Leu Lys Ser Asp Glu Asn Lys Glu Asn
 275 280 285

Ser Phe Ser Ala Asp His Val Thr Thr Ala Val Glu Lys Ser Lys Glu
 290 295 300

Ser Gln Val Thr Ala Asp Asp Leu Glu Glu Lys Ala Lys Ala Glu
 305 310 315 320

Leu Ile Met Asp Asp Asp Arg Thr Val Asp Pro Leu Leu Ser Lys Ser
 325 330 335

Gln Ser Ile Leu Ile Ser Thr Ser Ala Thr Ala Ser Ser Lys Lys Thr
 340 345 350

Ile Glu Asp Arg Asn Ile Lys Asn Lys Ser Thr Asn Asn Arg Ala
 355 360 365

Ser Ser Ala Ser Ala Arg Leu Met Thr Ser Glu Phe Leu Lys Lys Ser
 370 375 380

Ser Ser Lys Arg Arg Thr Pro Ser Thr Thr Ser Ser His Tyr Leu
 385 390 395 400

Gly Thr Leu Lys Val Leu Asp Gln Lys Pro Ser Gln Lys Gln Ser Ile
 405 410 415

Glu Pro

<210> 53

<211> 303

<212> PRT

<213> artificial sequence

<220>

<223> hASAP mutant 1-303

<400> 53

Met Ser Asp Glu Val Phe Ser Thr Thr Leu Ala Tyr Thr Lys Ser Pro
 1 5 10 15

s644PCT88.ST25

Lys Val Thr Lys Arg Thr Thr Phe Gln Asp Glu Leu Ile Arg Ala Ile
20 25 30

Thr Ala Arg Ser Ala Arg Gln Arg Ser Ser Glu Tyr Ser Asp Asp Phe
35 40 45

Asp Ser Asp Glu Ile Val Ser Leu Gly Asp Phe Ser Asp Thr Ser Ala
50 55 60

Asp Glu Asn Ser Val Asn Lys Lys Met Asn Asp Phe His Ile Ser Asp
65 70 75 80

Asp Glu Glu Lys Asn Pro Ser Lys Leu Leu Phe Leu Lys Thr Asn Lys
85 90 95

Ser Asn Gly Asn Ile Thr Lys Asp Glu Pro Val Cys Ala Ile Lys Asn
100 105 110

Glu Glu Glu Met Ala Pro Asp Gly Cys Glu Asp Ile Val Val Lys Ser
115 120 125

Phe Ser Glu Ser Gln Asn Lys Asp Glu Glu Phe Glu Lys Asp Lys Ile
130 135 140

Lys Met Lys Pro Lys Pro Arg Ile Leu Ser Ile Lys Ser Thr Ser Ser
145 150 155 160

Ala Glu Asn Asn Ser Leu Asp Thr Asp Asp His Phe Lys Pro Ser Pro
165 170 175

Trp Pro Arg Ser Met Leu Lys Lys Ser His Met Glu Glu Lys Asp
180 185 190

Gly Leu Glu Asp Lys Glu Thr Ala Leu Ser Glu Glu Leu Glu Leu His
195 200 205

Ser Ala Pro Ser Ser Leu Pro Thr Pro Asn Gly Ile Gln Leu Glu Ala
210 215 220

Glu Lys Lys Ala Phe Ser Glu Asn Leu Asp Pro Glu Asp Ser Cys Leu
225 230 235 240

Thr Ser Leu Ala Ser Ser Ser Leu Lys Gln Ile Leu Gly Asp Ser Phe
245 250 255

Ser Pro Gly Ser Glu Gly Asn Ala Ser Gly Lys Asp Pro Asn Glu Glu
260 265 270

Ile Thr Glu Asn His Asn Ser Leu Lys Ser Asp Glu Asn Lys Glu Asn
275 280 285

Ser Phe Ser Ala Asp His Val Thr Thr Ala Val Glu Lys Ser Lys
290 295 300